

**TESIS**

**DIAGNOSA KESEHATAN MENTAL MENGGUNAKAN K-NEAREST  
NEIGHBOR DENGAN WHALE OPTIMIZATION ALGORITHM**



Disusun oleh:

**Nama : Firza Septian**  
**NIM : 21.51.2110**  
**Konsentrasi : Digital Transformation Intelligence**

**PROGRAM STUDI S2 TEKNIK INFORMATIKA**  
**PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA**  
**YOGYAKARTA**

**2023**

**TESIS**

**DIAGNOSA KESEHATAN MENTAL MENGGUNAKAN K-NEAREST  
NEIGHBOR DENGAN WHALE OPTIMIZATION ALGORITHM**

**MENTAL HEALTH DIAGNOSIS USING K-NEAREST NEIGHBOR  
WITH WHALE OPTIMIZATION ALGORITHM**

Diajukan untuk memenuhi salah satu syarat memperoleh derajat Magister



Disusun oleh:

Nama : Firza Septian  
NIM : 21.51.2110  
Konsentrasi : Digital Transformation Intelligence

**PROGRAM STUDI S2 TEKNIK INFORMATIKA  
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2023**

**HALAMAN PENGESAHAN**

**DIAGNOSA KESEHATAN MENTAL MENGGUNAKAN K-NEAREST NEIGHBOR  
DENGAN WHALE OPTIMIZATION ALGORITHM**

**MENTAL HEALTH DIAGNOSIS USING K-NEAREST NEIGHBOR WITH WHALE  
OPTIMIZATION ALGORITHM**

Dipersiapkan dan Disusun oleh

**Firza Septian**

**21.51.2110**

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tesis  
Program Studi S2 Teknik Informatika  
Program Pascasarjana Universitas AMIKOM Yogyakarta  
pada hari Kamis, 08 Juni 2023

Tesis ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Magister Komputer

Yogyakarta, 08 Juni 2023

**Rektor**

**Prof. Dr. M. Suvanto, M.M.**  
**NIK. 190302001**

**HALAMAN PERSETUJUAN**  
**DIAGNOSA KESEHATAN MENTAL MENGGUNAKAN K-NEAREST NEIGHBOR**  
**DENGAN WHALE OPTIMIZATION ALGORITHM**

**MENTAL HEALTH DIAGNOSIS USING K-NEAREST NEIGHBOR**  
**WITH WHALE OPTIMIZATION ALGORITHM**

Dipersiapkan dan Disusun oleh

**Firza Septian**

**21.51.2110**

Telah Diujikan dalam Sidang Ujian Tesis  
Program Studi S2 Teknik Informatika  
Program Pascasarjana Universitas AMIKOM Yogyakarta  
pada hari Kamis, 08 Juni 2023

**Pembimbing Utama**

Prof. Dr. Kusriani, M.Kom.  
NIK. 190302106

**Anggota Tim Penguji**

Dr. Arief Setyanto, S.Si., M.T.  
NIK. 190302036

**Pembimbing Pendamping**

Tonny Hidayat, M.Kom., Ph.D.  
NIK. 190302182

Hanafi, S.Kom., M.Eng., Ph.D.  
NIK. 190302024

Prof. Dr. Kusriani, M.Kom.  
NIK. 190302106

Tesis ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Magister Komputer  
Yogyakarta, 08 Juni 2023

**Direktur Program Pascasarjana**

Prof. Dr. Kusriani, M.Kom.  
NIK. 190302106

## HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : Firza Septian  
NIM : 21.51.2110  
Konsentrasi : Digital Transformation Intelligence

Menyatakan bahwa Tesis dengan judul berikut:  
**DIAGNOSA KESEHATAN MENTAL MENGGUNAKAN K-NEAREST  
NEIGHBOR DENGAN WHALE OPTIMIZATION ALGORITHM**

Dosen Pembimbing Utama : Prof. Dr. Kusriani, M.Kom.  
Dosen Pembimbing Pendamping : Tonny Hidayat, M.Kom., Ph.D.

1. Karya tulis ini adalah benar-benar ASLI dan BELUMPERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi

Yogyakarta, 08 Juni 2023

Yang Menyatakan,



Firza Septian

## HALAMAN PERSEMBAHAN

Alhamdulillah, dengan rahmat dan karunia-Nya, penulis ingin mengungkapkan rasa syukur dengan menyampaikan rasa suka kepada

Tuhanku sampai kapanpun itu, **Allah SWT**

*Semoga persembahan ini menjadi bentuk kepatuhan kami kepada-Nya, dan semoga kami senantiasa mendapatkan petunjuk dan rahmat-Nya.*

**Kedua Orang Tuaku, Trisman dan Irianah**

*Terima kasih kepada ayah dan ibu kami, Trisman dan Irianah, yang telah memberi kami kasih, dukungan, dan arahan sepanjang perjalanan hidup kami. Kepedulian, dukungan, dan kasih sayang kalian adalah berkah yang tak ternilai.*

**Istriku Tersayang, Yossy Aprianita**

*Kepada istri tercinta, Yossy Aprianita, terima kasih telah membawa sinar ceria dalam hidupku. Cintamu, dukungan, dan keceriaanmu menjadi pelipur lara dalam setiap detik kehidupan kami. Semoga hubungan kita tetap diberkahi oleh Allah dan terus penuh dengan tawa dan canda.*

**Dedek Bayi, Said atau Saidah**

*Kamu sekarang masih 5 bulan nak, papi berkeinginan tuk memberikan engkau nama Said atau Saidah yang artinya Bahagia. Do'a papi engkau Bahagia dunia akhirat nak. Sebelum lahir saja, engkau sudah menjadi sinar kebahagiaan yang membawa tawa dan senyuman di antara kami. Semoga engkau tumbuh menjadi anak yang saleh, pintar, dan penuh keceriaan, dan senantiasa menjadi anugerah dari Allah SWT.*

**Kakak Sekeluarga, Engga Akbar, Kristi Okta Maria, Aqsha, Afnan, & Abdul**

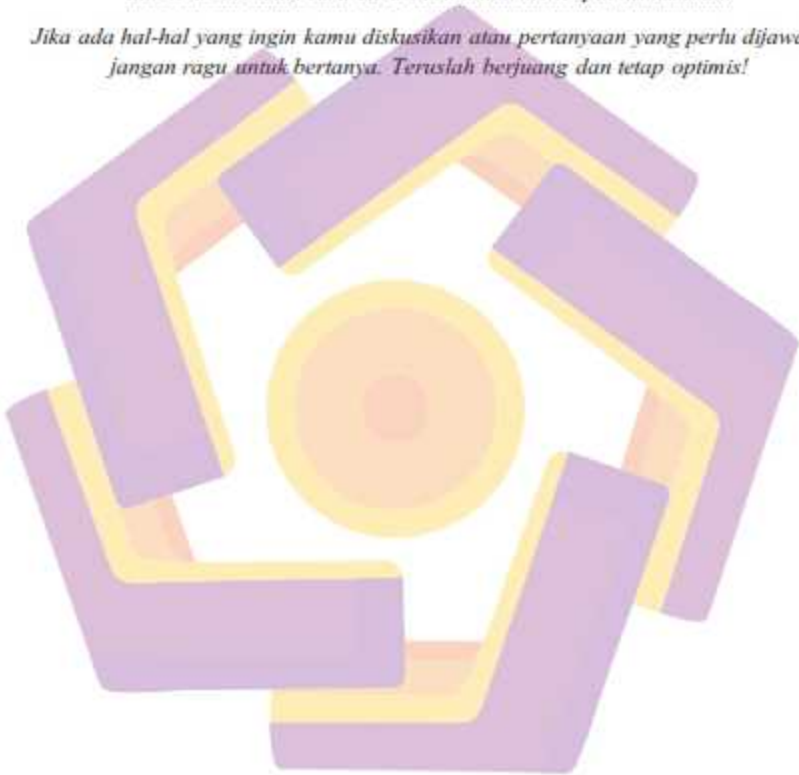
*Kepada kakakku dan keluargamu, Engga Akbar, terima kasih telah menjadi sumber inspirasi dan kebahagiaan bagi kami. Semoga keluarga kita tetap bersatu dalam cinta dan keceriaan.*

Dengan penuh rasa syukur, kami mengakhiri halaman persembahan ini. Semoga kebahagiaan, cinta, dan ceria senantiasa mengalir dalam setiap langkah hidup kami.

## HALAMAN MOTTO

*Semangat untuk terus meningkatkan diri dan mencapai target mingguanmu!  
Ingatlah untuk tetap konsisten, fokus, dan berusaha dengan sungguh-sungguh.  
Jangan ragu untuk belajar dari pengalamanmu, beradaptasi dengan perubahan,  
dan terus mencari cara-cara baru untuk mencapai kesuksesan.*

*Jika ada hal-hal yang ingin kamu diskusikan atau pertanyaan yang perlu dijawab,  
jangan ragu untuk bertanya. Teruslah berjuang dan tetap optimis!*



## KATA PENGANTAR

Kepada Tim Dosen Pembimbing, Tim Dosen Penguji, serta semua pihak yang terlibat dalam penyelesaian Tesis ini, saya ingin menyampaikan ucapan terima kasih yang sebesar-besarnya. Ucapan terima kasih ini saya sampaikan dengan penuh rasa hormat dan tulus, tanpa nama-nama yang tidak terkait.

Pertama-tama, kepada Pembimbing Utama, Prof. Dr. Kusri, M.Kom., saya mengucapkan terima kasih atas bimbingan, arahan, dan ilmu yang berharga selama proses penulisan Tesis ini. Tanpa bimbingan Anda, pencapaian ini tidak akan terwujud.

Kepada Pembimbing Pendamping, Tonny Hidayat M.Kom., Ph.D., saya juga berterima kasih atas dukungan dan pandangan yang berharga dalam penyusunan Tesis ini. Kontribusi Anda sangat berarti bagi perkembangan penelitian ini.

Tak lupa, saya ingin mengucapkan terima kasih kepada Ayah, Ir. Trisman, dan Ibuku, Iriah, atas doa, dukungan moral, dan semangat yang selalu Anda berikan dalam perjalanan ini. Semua ini tidak akan berhasil tanpa kasih sayang dan doa dari orang tua.


Saya juga ingin menyampaikan rasa terima kasih kepada Kakakku, Engga Akbar, S.Pd.T., dan Istrinya, Kristi Okta Maria, S.Pd., serta anak-anaknya, Aqsha, Afnan, dan Abdul, atas dukungan dan pengertian selama penulisan Tesis ini.



Kepada Istriku, Yossy Aprianita, A.Md.Kom., terima kasih atas kesabaran, dukungan, dan cinta yang selalu Anda berikan. Anda adalah sumber inspirasi terbesar dalam hidup saya.

Terakhir, namun tidak kalah penting, kepada calon anakku, doa dan harapan saya untukmu tidak terhingga. Semoga Tesis ini menjadi bukti bahwa kita selalu berjuang untuk menjadi lebih baik demi masa depan yang cerah bersama.

Segegap ucapan terima kasih ini saya sampaikan dengan rendah hati. Semoga semua pihak yang telah memberikan kontribusi dalam penyelesaian Tesis ini senantiasa diberkati dalam segala hal. Terima kasih.



Yogyakarta, 08 Juni 2023

Penulis

## DAFTAR ISI

|  |      |
|--|------|
| HALAMAN JUDUL.....                     | ii   |
| HALAMAN PENGESAHAN.....                | iii  |
| HALAMAN PERSETUJUAN.....               | iv   |
| HALAMAN PERNYATAAN KEASLIAN TESIS..... | v    |
| HALAMAN PERSEMBAHAN.....               | vi   |
| HALAMAN MOTTO.....                     | vii  |
| KATA PENGANTAR.....                    | viii |
| DAFTAR ISI.....                        | x    |
| DAFTAR TABEL.....                      | xiii |
| DAFTAR GAMBAR.....                     | xiv  |
| INTISARI.....                          | xv   |
| <i>ABSTRACT</i> .....                  | xvi  |
| BAB I PENDAHULUAN.....                 | 1    |
| 1.1. Latar Belakang Masalah.....       | 1    |
| 1.2. Rumusan Masalah.....              | 10   |
| 1.3. Batasan Masalah.....              | 10   |
| 1.4. Tujuan Penelitian.....            | 11   |
| 1.5. Manfaat Penelitian.....           | 11   |
| BAB II TINJAUAN PUSTAKA.....           | 13   |
| 2.1. Tinjauan Pustaka.....             | 13   |
| 2.2. Keaslian Penelitian.....          | 21   |

|   |           |
|---|-----------|
| 2.3. Landasan Teori.....                              | 27        |
| 2.3.1 Mental Health .....                             | 27        |
| 2.3.2 WOA.....  | 30        |
| 2.3.3 Algoritma Pembentuk Deteksi Mental Health ..... | 34        |
| <b>BAB III METODE PENELITIAN.....</b>                 | <b>42</b> |
| 3.1. Jenis, Sifat, dan Pendekatan Penelitian.....     | 42        |
| 3.2. Metode Pengumpulan Data.....                     | 43        |
| 3.3. Metode Analisis Data.....                        | 43        |
| 3.4. Alur Penelitian .....                            | 44        |
| <b>BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....</b>    | <b>46</b> |
| 4.1 Prosedur Baru.....                                | 46        |
| 4.1.1 Perumusan .....                                 | 46        |
| 4.1.2 Langkah Implementasi dalam Komputasi.....       | 51        |
| 4.2 Action Planning.....                              | 52        |
| 4.2.1 Pengumpulan Data.....                           | 52        |
| 4.2.2 Preprocessing Data .....                        | 56        |
| 4.3 Processing Data.....                              | 59        |
| 4.3.1 Google Colab.....                               | 59        |
| 4.3.2 Microsoft Excel .....                           | 65        |
| 4.3.3 Variabel yang digunakan.....                    | 71        |
| 4.3.4 KNN.....  | 83        |

|   |     |
|---|-----|
| 4.3.5 KNN WOA.....                              | 84  |
| 4.3.6 KNN-EEWOA.....                            | 85  |
| 4.4 Evaluating.....                             | 89  |
| 4.4.1 KNN.....                                  | 89  |
| 4.4.2 KNN WOA.....                              | 91  |
| 4.4.3 KNN Edge Enhancement WOA (KNN-EEWOA)..... | 93  |
| 4.4.4 Analisa Standar Deviasi.....              | 95  |
| 4.4.5 Comparison.....                           | 96  |
| BAB V PENUTUP.....                              | 99  |
| 5.1. Kesimpulan.....                            | 99  |
| 5.2. Saran.....                                 | 100 |
| DAFTAR PUSTAKA.....                             | 103 |

## DAFTAR TABEL

|   |    |
|---|----|
| Tabel 2.1 Mental Health Classification dengan 10 Pengklasifikasi..... | 16 |
| Tabel 2.2 Matriks literatur review dan posisi penelitian .....        | 21 |
| Tabel 4.1 Dataset Sebelum di Olah .....                               | 53 |
| Tabel 4.2 Data Preprocessing .....                                    | 58 |
| Tabel 4.3 Data Hasil Pengolahan .....                                 | 87 |
| Tabel 4.4 Beberapa Data Hasil Pengujian.....                          | 88 |
| Tabel 4.5 Perbandingan Hasil Evaluasi pada Google Colab.....          | 96 |
| Tabel 4.6 Hasil Comparasi pada Ms Excel .....                         | 97 |

## DAFTAR GAMBAR

|   |    |
|---|----|
| Gambar 1.1 Hasil Spesifikasi Penelitian Rajeshkumar & Kousalya .....              | 7  |
| Gambar 1.2 Hasil Sensitifitas Penelitian Rajeshkumar & Kousalya .....             | 7  |
| Gambar 2.1 Contoh Kasus KNN .....   | 35 |
| Gambar 3.1 Alur Penelitian .....  | 44 |
| Gambar 4.1 Penentuan Lintasan Batas Berdasarkan Intersection .....                | 48 |
| Gambar 4.2 KNN-WOA .....  | 49 |
| Gambar 4.3 KNN-EEWOA .....  | 50 |
| Gambar 4.4 Confusion Matrix data 1001-1050 pada KNN dengan $k = 4$ .....          | 66 |
| Gambar 4.5 Pemrosesan Data dengan KNN-WOA .....                                   | 67 |
| Gambar 4.6 Confusion Matrix data 1001-1050 pada KNN-WOA<br>dengan $k = 9$ .....   | 68 |
| Gambar 4.7 Confusion Matrix data 1001-1050 pada KNN-EEWOA<br>dengan $k = 9$ ..... | 69 |
| Gambar 4.8 Pemrosesan Data dengan KNN-EEWOA .....                                 | 70 |

## INTISARI

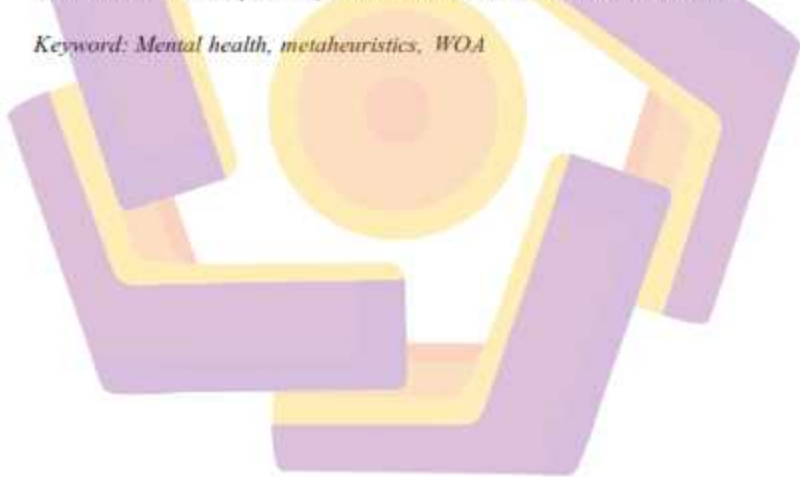
Diagnosis kesehatan mental dengan waktu secepat dan seakurat mungkin diperlukan mengingat banyak orang yang berlingkungan tingkat kebisingan tinggi serta berpotensi mengalami penyakit natural ini. Penulis mengajukan algoritma metaheuristik yang bersifat natural. KNN dan WOA dapat digunakan bersama-sama untuk melakukan diagnosis kesehatan mental. Penelitian yang berjudul Diagnosis Kesehatan Mental Menggunakan K-Nearest Neighbor dengan Whale Optimization Algorithm ini menjelaskan penggunaan kombinasi algoritma K-Nearest Neighbor (KNN) dengan Whale Optimization Algorithm (WOA). Dilakukan sedikit pengembangan pada bagian batas didalam WOA untuk membantu mengurangi jumlah iterasi yang musti dilakukan. Alhasil pengembangan ini diberi nama Edge Enchanment WOA (EEWOA). Metode eksperimental difokuskan pada tiga jenis algoritma: (1) KNN, (2) KNN-WOA, (3) KNN-EEWOA. Hasil membuktikan tingkat akurasi terbesar ke terkecil secara berurutan ialah KNN-EEWOA, KNN-WOA, KNN. Selain itu jumlah iterasi KNN-EEWOA lebih rendah dibandingkan dengan jumlah iterasi KNN-WOA. Kendati demikian, waktu pengukuran tercepat secara berurutan ialah KNN, KNN-WOA, dan KNN-EEWOA.

Kata kunci: Kesehatan mental, metaheuristik, WOA

## ABSTRACT

*Diagnosis of mental health in the quickest and most accurate time possible is necessary considering that many people live in high-noise environments and have the potential to experience this natural disease. The author proposes a natural metaheuristic algorithm. KNN and WOA can be used together to make a mental health diagnosis. The research entitled Mental Health Diagnosis Using the K-Nearest Neighbor with the Whale Optimization Algorithm explains the combined use of the K-Nearest Neighbor (KNN) algorithm with the Whale Optimization Algorithm (WOA). A little development was done on the boundaries in WOA to help reduce the number of iterations that must be done. As a result, this development was named Edge Enchantment WOA (EEWOA). The experimental method is focused on three types of algorithms: (1) KNN, (2) KNN-WOA, (3) KNN-EEWOA. The results prove the highest to the smallest level of accuracy in sequence, namely KNN-EEWOA, KNN-WOA, KNN. In addition, the number of KNN-EEWOA iterations is lower than the number of KNN-WOA iterations. However, the fastest measurement time sequentially is KNN, KNN-WOA, and KNN-EEWOA.*

*Keyword: Mental health, metaheuristics, WOA*





# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Saat ini, mayoritas orang tinggal di wilayah metropolitan. Gupta et al. (2018) menyatakan bahwa berbagai faktor seperti kebisingan berdampak negatif pada kesehatan mental orang-orang dari segala usia, bahkan anak-anak yang belum lahir, hal ini terjadi di sebuah kota di India utara. Pengecekan kesehatan mental diperlukan karenanya dengan efisiensi waktu yang perlu di optimalkan.

Memang benar bahwa tinggal di wilayah metropolitan dapat memiliki dampak negatif pada kesehatan mental seseorang, dan hal ini disebabkan oleh berbagai faktor seperti kebisingan, polusi udara, kepadatan penduduk, dan tekanan sosial yang tinggi. Studi seperti yang dilakukan oleh Gupta et al. (2018) menunjukkan bahwa faktor-faktor tersebut dapat berdampak buruk pada kesehatan mental orang-orang dari segala usia.

Untuk memeriksa kesehatan mental, ada beberapa pilihan yang bisa dipertimbangkan, tergantung pada situasi dan kebutuhan individu. Beberapa pilihan tersebut antara lain adalah: (1) Konsultasi dengan dokter atau psikolog: Jika seseorang merasa mengalami masalah kesehatan mental, maka konsultasi dengan dokter atau psikolog dapat menjadi pilihan yang baik. Mereka dapat membantu melakukan penilaian kesehatan mental dan memberikan saran dan pengobatan yang sesuai. (2) Mengikuti kelompok dukungan: Kelompok dukungan merupakan forum di mana individu dapat berbagi pengalaman dan mendapat dukungan dari orang lain

yang mengalami hal serupa. Kelompok dukungan dapat dilakukan secara online atau offline. (3) Mengikuti program kesehatan mental online: Ada banyak program kesehatan mental online yang tersedia di internet, seperti aplikasi kesehatan mental, forum diskusi online, dan webinar. Program ini dapat membantu individu memperbaiki kesehatan mental mereka secara mandiri. Dalam program Kesehatan mental online, tentu pengecekan awal atau diagnosis awal dilakukan secara online dan dianalisa menggunakan algoritma-algoritma tertentu (Latifah et al., 2023). Tingkat akurasi yang cukup tinggi dapat diraih dengan memanfaatkan kombinasi antar algoritma yang telah ada sebelumnya.

Untuk memaksimalkan efisiensi waktu, individu dapat mempertimbangkan beberapa pilihan di atas yang cocok dengan jadwal mereka. Konsultasi dengan dokter atau psikolog biasanya membutuhkan waktu yang lebih lama dan memerlukan jadwal yang fleksibel, sedangkan mengikuti program kesehatan mental online atau kelompok dukungan dapat dilakukan sesuai dengan waktu luang individu (Ouriaghli et al., 2019).

Sebenarnya pada tahun 2017, IBM mengambil langkah untuk mengatasi permasalahan kesehatan mental. IBM meramalkan bahwa Artificial Intelligence (AI) akan mentransformasi implementator dalam dunia kesehatan mental selama lima tahun ke depan khususnya perihal diagnostik. Perawatan oleh dokter akan menjadi lebih baik, lebih akurat dalam memprediksi, memantau, dan melacak kondisi pasien (IBM, 2017).

Teknologi AI dapat membantu dokter dan profesional kesehatan mental dalam melakukan penilaian dan diagnosis yang lebih akurat dan cepat. AI dapat

memproses data dan informasi dari berbagai sumber, seperti rekam medis elektronik, tes psikologis, dan hasil pemantauan fisik dan perilaku pasien (Graham et al., 2019). Dengan menggunakan teknologi AI, dokter dapat membuat diagnosis yang lebih tepat dan efektif, serta memberikan perawatan yang lebih terfokus dan terpersonalisasi.

Selain itu, teknologi AI juga dapat membantu dalam memantau dan melacak kondisi pasien secara terus-menerus. AI dapat mengumpulkan dan menganalisis data dari berbagai sensor dan perangkat, seperti monitor detak jantung, monitor tidur, dan sensor aktivitas fisik (Gomes et al., 2023). Dengan demikian, dokter dapat melihat perubahan kondisi pasien secara real-time dan dapat segera mengambil tindakan yang diperlukan.

Sederhananya, "AI adalah bidang ilmu komputer yang mencakup robotika, pemrosesan bahasa alami, pemrosesan ucapan, pembelajaran mesin [algoritma], dan pengambilan keputusan otomatis terkait" (Hodgson, Berry, Wearne, & Ellis, 2018). Psikiatri juga tampaknya percaya pada kemampuan teknologi semacam itu untuk meningkatkan taraf kehidupan manusia (Wykes, 2019), yakni tentang intervensi kesehatan mental berbasis AI (Bhugra, 2017). Psikiatri dan profesional kesehatan mental lainnya juga percaya pada kemampuan teknologi AI untuk meningkatkan taraf kehidupan manusia dengan memberikan intervensi kesehatan mental yang lebih baik dan efektif. Dalam beberapa tahun terakhir, telah ada banyak penelitian dan pengembangan terkait dengan aplikasi teknologi AI dalam bidang kesehatan mental. Investigasi skala besar mental health pada data pasien itu masih jarang dilakukan (Jami, 2021) meskipun beberapa faktor utama dalam

mengidentifikasi mental health telah dipelajari oleh Morales dengan network analysis, node-node khusus pada risk factors diidentifikasi dengan node yang paling sentral seperti contohnya kecemasan, tekanan yang dirasakan, tidak adanya energi, dan lain-lain (Morales, 2017). Algoritma untuk membantu mengklasifikasikan skala besar diperlukan dalam hal ini. Terutama algoritma dengan tingkat spesifiktas yang tinggi.

Dalam penerapan teknologi AI untuk diagnosa dan perawatan kesehatan mental, algoritma yang dapat membantu mengklasifikasikan skala besar data sangat diperlukan (Lee et al., 2021). Algoritma ini harus memiliki tingkat spesifiktas yang tinggi agar dapat membedakan gejala dan kondisi kesehatan mental yang berbeda secara akurat. Algoritma pembelajaran mesin dapat dilatih dengan menggunakan data rekam medis, tes psikologis, dan survei pasien untuk mengidentifikasi pola dan korelasi antara gejala dan kondisi kesehatan mental yang berbeda (Bhugra, 2017).

Dengan menggunakan teknik machine learning, algoritma dapat terus belajar dan berkembang seiring dengan bertambahnya data dan informasi yang diterima. Hal ini dapat membantu meningkatkan tingkat spesifiktas algoritma dalam mengklasifikasikan data kesehatan mental secara akurat (Graham et al., 2019).

Namun, penting juga untuk memastikan bahwa algoritma yang dikembangkan didasarkan pada data yang akurat dan terpercaya, serta diperiksa secara teratur oleh dokter dan profesional kesehatan mental untuk memastikan keandalan dan validitas hasil yang diberikan.

Whale Optimization Algorithm (WOA) merupakan algoritma berbasis metaheuristik yang meniru strategi perburuan bubble-net paus bungkuk. Dalam algoritma WOA, populasi solusi digambarkan sebagai sekumpulan paus yang bergerak menuju target atau mangsa (solusi optimal). Setiap paus dalam populasi memiliki posisi yang mewakili solusi dan kecepatan yang menentukan arah pergerakan paus dalam mencari solusi optimal. Algoritma WOA terdiri dari tiga tahap, yaitu eksplorasi, eksploitasi, dan update posisi paus. Tahap eksplorasi dilakukan dengan menggerakkan paus secara acak untuk mengeksplorasi area pencarian solusi. Tahap eksploitasi dilakukan dengan menggerakkan paus ke arah solusi terbaik yang ditemukan pada tahap eksplorasi. Tahap update posisi paus dilakukan dengan memperbarui posisi paus berdasarkan solusi terbaik yang ditemukan pada tahap eksploitasi (Mirjalili, 2016).

Algoritma WOA telah digunakan dalam berbagai aplikasi optimasi, termasuk dalam bidang kesehatan mental. Beberapa penelitian telah menunjukkan bahwa algoritma WOA dapat digunakan untuk mengoptimalkan parameter dalam model prediksi kesehatan mental, seperti model prediksi depresi dan kecemasan. Dengan menggunakan algoritma WOA, peneliti dapat mencapai solusi yang lebih akurat dan efisien dalam waktu yang lebih singkat. Rajeshkumar dan Kousalya menyatakan bahwa WOA dengan algoritma backpropagation neural network mampu mengoptimalkan masalah klasifikasi tipe diabetes (Rajeshkumar dan Kousalya, 2017).

Point penting disini bahwa WOA merupakan algoritma optimasi, jadi masih dibutuhkan algoritma utama untuk menjalankannya. Berbagai macam kombinasi

kemungkinan dapat dilakukan. Salah satu kombinasi yang memungkinkan untuk dilakukan ialah kombinasi antara WOA dengan K Nearest Neighbor atau bisa disebut dengan KNN-WOA yang dilakukan oleh Hajjhashemi dan publish pada tahun 2019.

KNN-WOA merupakan salah satu metode klasifikasi yang digunakan dalam diagnostik kesehatan mental dengan menggunakan kombinasi algoritma k-nearest neighbor (KNN) dan whale optimization algorithm (WOA). Metode ini bertujuan untuk memprediksi diagnosis kesehatan mental berdasarkan data pasien.

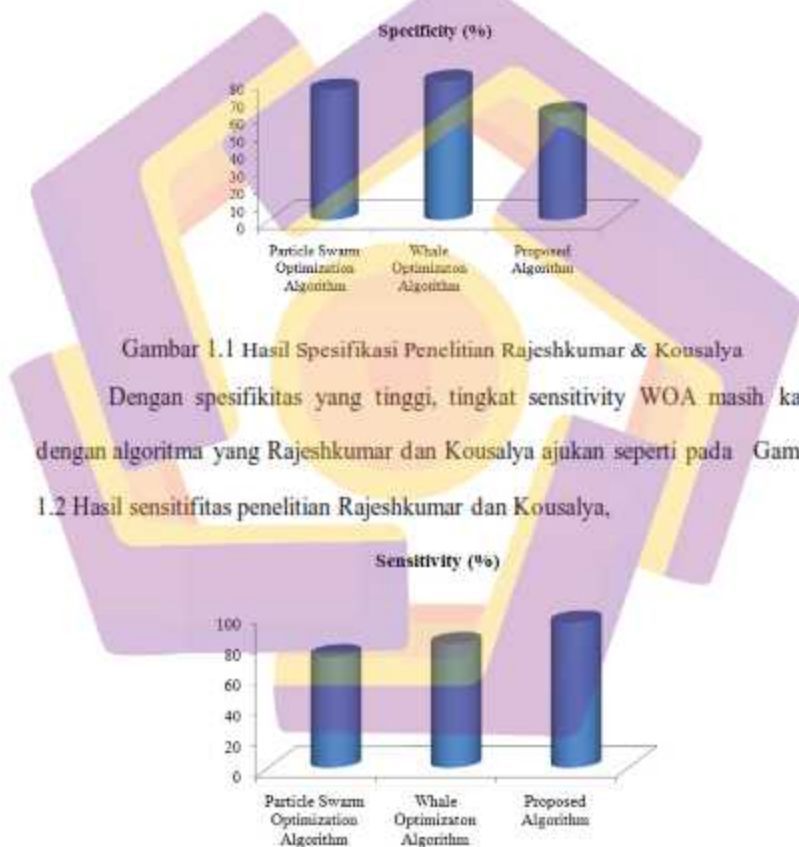
Algoritma KNN digunakan untuk mengklasifikasikan data pasien berdasarkan kategori diagnosis kesehatan mental yang sudah diketahui. Dalam algoritma KNN, setiap data pasien akan diklasifikasikan ke dalam kategori diagnosis berdasarkan kategori diagnosis yang paling banyak muncul pada K tetangga terdekat dari data pasien tersebut.

Algoritma WOA digunakan untuk mengoptimalkan parameter pada algoritma KNN, seperti nilai K (jumlah tetangga terdekat yang akan dipertimbangkan) dan bobot untuk masing-masing tetangga terdekat. Dengan menggunakan algoritma WOA, nilai parameter pada algoritma KNN dapat diatur secara optimal untuk menghasilkan akurasi klasifikasi yang lebih tinggi.

Dalam penelitian diagnostik kesehatan mental, KNN-WOA telah digunakan untuk memprediksi diagnosis berbagai jenis gangguan kesehatan mental, seperti depresi, kecemasan, dan stres. Hasil penelitian menunjukkan bahwa KNN-WOA dapat menghasilkan akurasi klasifikasi yang lebih tinggi dibandingkan dengan

menggunakan algoritma KNN tanpa optimasi menggunakan algoritma WOA (Hajjhashemi et al., 2019).

Berbarengan dengan algoritma natural lainnya seperti Particle Swarm Optimization, WOA menunjukkan kelebihanya dibidang spesifikitasnya seperti pada Gambar 1.1 Hasil spesifikitas penelitian Rajeshkumar dan Kousalya.



Gambar 1.1 Hasil Spesifikasi Penelitian Rajeshkumar & Kousalya

Dengan spesifikitas yang tinggi, tingkat sensitivity WOA masih kalah dengan algoritma yang Rajeshkumar dan Kousalya ajukan seperti pada Gambar 1.2 Hasil sensitifitas penelitian Rajeshkumar dan Kousalya,

Gambar 1.2 Hasil Sensitifitas Penelitian Rajeshkumar & Kousalya

Penerapan WOA pada klasifikasi diagnostic diabetes tergolong berhasil ketika dibandingkan dengan algoritma natural Particle Swarm Optimization

Algorithm (PSOA). Implementasi WOA pada klasifikasi mental health diperlukan pada data skala besar mental health.

Kemudian alasan pendukung seperti beberapa jurnal di *Journal of Mental Health* belum ada yang membahas mengenai prediksi mental health dengan menggunakan algoritma natural tipe whale optimization algorithm. Setiap penelitian dan perkembangan teknologi membutuhkan waktu dan usaha yang cukup untuk menghasilkan hasil yang valid dan signifikan. Meskipun WOA telah terbukti efektif dalam aplikasi pada berbagai bidang, namun penggunaannya dalam prediksi kesehatan mental masih relatif baru dan perlu penelitian yang lebih lanjut.

Jurnal-jurnal di *Journal of Mental Health* belum mengulas mengenai prediksi kesehatan mental menggunakan algoritma natural tipe WOA mungkin karena penelitian tersebut masih dalam tahap awal dan belum mencapai hasil yang signifikan. Namun, ini bukan berarti bahwa WOA tidak memiliki potensi untuk digunakan dalam prediksi kesehatan mental di masa depan.

Dilanjutkan dengan alasan pendukung lainnya yakni, bahkan cloud computing sekelas AWS Amazon belum menyatakan bahwa mereka menerima pesanan berupa layanan prediksi mental health. Pada laman resminya, AWS Amazon mengatakan bahwa mereka dengan platform dan layanan pembelajaran mesin terlengkap untuk industri kesehatan, machine learning AWS memberikan perusahaan berupa perawatan kesehatan serta artificial intelligence yang mereka butuhkan untuk memahami pola dari berbagai data pasien mereka. (AWS Amazon, 2022).



Hal ini menunjukkan bahwa AWS Amazon juga menyadari potensi dan pentingnya teknologi machine learning dalam bidang kesehatan, termasuk kesehatan mental. AWS Amazon bahkan telah bekerja sama dengan beberapa perusahaan kesehatan dan peneliti untuk mengembangkan solusi kesehatan berbasis machine learning yang dapat membantu dalam diagnosis dan perawatan pasien.

Dalam hal ini, penelitian dan pengembangan teknologi prediksi kesehatan mental menggunakan algoritma seperti WOA adalah langkah awal yang penting untuk mengembangkan solusi kesehatan mental berbasis teknologi. Dalam jangka panjang, solusi ini dapat membantu memperbaiki dan meningkatkan perawatan kesehatan mental yang tersedia.

Jadi algoritma metaheuristic yang bersifat natural seperti KNN-WOA diperlukan untuk mendiagnosis mental health untuk membantu peran psikiatris dalam pekerjaannya agar mental health seseorang tetap terjaga aman dan tidak melebihi batas amannya. Pengembangan KNN-WOA dirasa akan sedikit membantu dunia diagnosis. Oleh sebab itu, prosedur yang dapat dikembangkan untuk meningkatkan akurasi diperlukan. Selain itu pengukuran tingkat akurasi antara KNN, KNN-WOA, dan prosedur baru diperlukan agar dapat mengevaluasi kinerja dari setiap model tersebut. Evaluasi dilakukan dengan memfokuskan diri pada analisa tingkat akurasi yang telah dicapai pada masing-masing algoritma yang telah diajukan.

## 1.2. Rumusan Masalah

Berdasarkan implementasi algoritma natural pada klasifikasi diagnosis mental health skala besar belum dilakukan, WOA dapat menjadi solusinya. Kemudian algoritma dengan tingkat spesifiktas tinggi diperlukan dengan tidak beracuan pada time processing. Keduanya digunakan untuk menentukan rumusan masalah berupa:

- a. Algoritma baru apakah yang dapat diajukan untuk meningkatkan tingkat akurasi pada WOA?
- b. Berapa tingkat akurasi atau tingkat spesifiktas KNN, KNN-WOA, serta algoritma baru yang diajukan pada diagnostik mental health?

## 1.3. Batasan Masalah

Penelitian ini dilakukan dengan menyatakan berbagai batasan masalah seperti:

- a. Fokus pada pengembangan algoritma WOA untuk meningkatkan tingkat akurasi pada WOA, dengan mengoptimalkan parameter-parameter yang terlibat dalam algoritma, seperti batas awal bubble pada WOA sehingga intersection tidak terjadi.
- b. Mengukur tingkat akurasi atau tingkat spesifiktas KNN, KNN-WOA, dan algoritma baru yang diajukan pada diagnostik mental health, dengan menggunakan data sampel dan representatif dari populasi yang relevan.

#### 1.4. Tujuan Penelitian

Tujuan dilakukannya penelitian ini secara spesifik ialah:

- a. Untuk mengembangkan sebuah algoritma baru yang dapat meningkatkan tingkat akurasi pada algoritma WOA dalam melakukan prediksi kesehatan mental.
- b. Selain itu, penelitian ini juga bertujuan untuk membandingkan tingkat akurasi atau tingkat spesifiktas dari KNN, KNN-WOA, dan algoritma baru yang diajukan pada diagnostik kesehatan mental, dengan menggunakan data sampel yang cukup besar dan representatif dari populasi yang relevan.

Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan teknologi untuk mendukung bidang kesehatan mental dan memberikan informasi yang lebih akurat bagi para profesional kesehatan dalam melakukan diagnosis gangguan kesehatan mental.

#### 1.5. Manfaat Penelitian

Dengan dilakukannya penelitian ini, kedepannya pembaca dapat memperoleh berbagai manfaat seperti:

- a. Memberikan kontribusi pada pengembangan teknologi kesehatan mental: Dengan mengembangkan algoritma baru untuk meningkatkan akurasi pada algoritma WOA, penelitian ini dapat memberikan kontribusi pada pengembangan teknologi yang dapat membantu para profesional kesehatan dalam melakukan diagnosis dalam sesi perawatan pasien dengan gangguan kesehatan mental.

- b. Memberikan informasi yang lebih akurat dalam diagnosis kesehatan mental: Dengan membandingkan tingkat akurasi dari KNN, KNN-WOA, dan algoritma baru yang diajukan pada diagnostik kesehatan mental, penelitian ini dapat memberikan informasi yang lebih akurat bagi para profesional kesehatan dalam melakukan diagnosis dalam perawatan pasien.
- c. Meningkatkan efisiensi dan efektivitas perawatan. Dengan menggunakan teknologi yang lebih canggih dan akurat, penelitian ini dapat membantu para profesional kesehatan dalam meningkatkan efisiensi dan efektivitas perawatan pasien dengan gangguan kesehatan mental.
- d. Memberikan kontribusi pada peningkatan kualitas hidup pasien. Dengan memberikan informasi yang lebih akurat yang lebih efektif, penelitian ini dapat memberikan kontribusi pada peningkatan kualitas hidup pasien dengan gangguan kesehatan mental.

## BAB II

### TINJAUAN PUSTAKA

#### 2.1. Tinjauan Pustaka

Beberapa penelitian telah dilakukan berkaitan dengan implementasi WOA pada dataset tertentu maupun implementasi algoritma lainnya pada mental health. Tinjauan pustaka dan acuan pada penelitian ini diantaranya ialah:

Kemudian penelitian berkaitan WOA di dunia medic yang dilakukan oleh Vaiyapuri & Alaskar (2020) berfokus ke implementasinya ke CT dan MR images. Pekerjaan mengisi celah dalam literatur dengan memeriksa untuk pertama kalinya berbagai algoritma kecerdasan untuk segmentasi gambar medis tanpa pengawasan, termasuk fuzzy, algoritma genetika (GA), dan optimisasi segerombolan partikel (PSO). Ini akan menunjukkan keefektifan metode yang disarankan. Gambar CT dan MR medis dievaluasi menggunakan kesamaan fitur (FSIM), dadu (DC), dan fitur kelebihan (FOM). Hasil dari eksperimen menunjukkan betapa unggulnya pendekatan yang disarankan dibandingkan algoritma intelijen alternatif. Untuk menunjukkan pentingnya pendekatan yang disarankan dalam memilih solusi terbaik dan menunjukkan hasil segmentasi yang menjanjikan terhadap dukungan diagnostik untuk ahli radiologi, investigasi statistik dengan analisis ANOVA telah diselesaikan. Penelitian yang dilakukan oleh Vaiyapuri & Alaskar (2020) mengisi celah dalam literatur dengan menguji berbagai algoritma kecerdasan untuk segmentasi gambar medis tanpa pengawasan, termasuk fuzzy, algoritma genetika (GA), dan optimisasi segerombolan partikel (PSO). Penelitian ini menunjukkan

keefektifan algoritma WOA dalam memilih solusi terbaik dan menunjukkan hasil segmentasi yang menjanjikan untuk dukungan diagnostik bagi ahli radiologi. Manfaat dari penelitian ini adalah memberikan kontribusi pada pengembangan teknologi di bidang kedokteran, khususnya dalam segmentasi gambar medis untuk diagnosis dan pengobatan. Selain itu, hasil penelitian ini dapat digunakan sebagai acuan untuk pengembangan algoritma segmentasi gambar medis tanpa pengawasan pada masa mendatang.

Penelitian terdahulu yang menggunakan WOA dilakukan oleh Shahraki et al. (2022) dengan case study berupa deteksi Covid-19. WOA yang digunakan ialah BE-WOA dimana BE singkatan dari Binary Enhanced. Algoritma ini dapat mendeteksi Covid-19 dengan tingkat akurasi sebesar 96.409%. Perlu dicatat bahwa penelitian ini tidak secara langsung berkaitan dengan deteksi kesehatan mental, namun penelitian ini memberikan pandangan mengenai efektivitas algoritma WOA dalam masalah klasifikasi dan deteksi. Oleh karena itu, hal ini dapat memberikan dasar untuk penelitian dalam meningkatkan akurasi pada masalah diagnostik kesehatan mental menggunakan algoritma KNN-WOA.

Bidang lain pada WOA yang dikerjakan oleh Wang et al. (2022) dengan algoritma Whale Optimization Support Vector Machine pada diagnostik Covid-19. Tingkat akurasi mencapai 72% dengan presisi mencapai 73%. Output dari setiap MSVM digunakan sebagai bukti independen untuk membangun distribusi reliabilitas fundamental pada indeks diagnostik ini. Pengklasifikasi indeks tunggal COVID-19 dibuat menggunakan MSVM yang dioptimalkan paus dan didasarkan pada gambar CT, deteksi indeks laboratorium rutin, sitokin serum, dan deteksi

kemokin. Penelitian yang dilakukan oleh Wang et al. (2022) ini menggunakan algoritma Whale Optimization Support Vector Machine (WOA-SVM) pada diagnostik Covid-19. Tujuannya adalah untuk mengembangkan sebuah indeks tunggal untuk mendeteksi Covid-19 yang dapat diandalkan dan akurat. WOA-SVM digunakan untuk mengoptimalkan parameter-parameter SVM, seperti kernel dan parameter C, sehingga meningkatkan akurasi dari model SVM. Selain itu, penelitian ini juga menggunakan data dari berbagai sumber, seperti gambar CT, deteksi indeks laboratorium rutin, sitokin serum, dan deteksi kemokin, untuk membangun indeks tunggal Covid-19 yang lebih akurat. Hasil penelitian menunjukkan bahwa indeks tunggal Covid-19 yang dibuat menggunakan WOA-SVM memiliki tingkat akurasi mencapai 72% dengan presisi mencapai 73%. Output dari setiap SVM juga digunakan sebagai bukti independen untuk membangun distribusi reliabilitas fundamental pada indeks diagnostik ini. Dalam bidang diagnostik Covid-19, penelitian ini membuktikan bahwa penggunaan WOA-SVM dapat meningkatkan akurasi model diagnostik, sehingga dapat membantu para ahli medis dalam melakukan deteksi Covid-19 yang lebih cepat dan akurat.

Studi yang dilakukan oleh Chung & Teo (2022) bertujuan untuk membandingkan kinerja berbagai algoritma klasifikasi dalam memprediksi status kesehatan mental individu berdasarkan data klinis. Studi ini melibatkan 309 partisipan dan 6 variabel prediktif. Sepuluh algoritma klasifikasi yang dibandingkan adalah Bayesian Network, Naive Bayes, Logistic Regression, Multiple Layer Perceptron, Sequential Minimal Optimisation, K-star, Random

Subspace, J48, Random Forest, dan Random Tree seperti yang ditunjukkan pada Tabel 2.1.

Tabel 2.1 Mental Health Classification dengan 10 Pengklasifikasi

| Klasifikator                    | Evaluation Matrix (%) |           |      |
|---------------------------------|-----------------------|-----------|------|
|                                 | Akurasi               | F-measure | AUC  |
| Bayesian network                | 79.8                  | 79.7      | 88.9 |
| Naive Bayes                     | 79.6                  | 79.4      | 85.3 |
| Logistic regression             | 72.4                  | 72.2      | 81.1 |
| Multiple layer perceptron       | 77.8                  | 77.8      | 85.0 |
| Sequential minimal optimisation | 75.3                  | 74.6      | 75.9 |
| K-star                          | 75.3                  | 75.3      | 81.4 |
| Random subspace                 | 87.5                  | 87.5      | 91.7 |
| J48                             | 87.8                  | 87.8      | 86.0 |
| Random forest                   | 89.0                  | 89.0      | 94.3 |
| Random tree                     | 85.1                  | 85.1      | 85.0 |

Penelitian pada mental health ini dilakukan dengan berbagai jenis klasifikator, namun belum digunakan natural algorithm di dalamnya. Terdapat 30 paper yang digunakan dalam analisa kasus ini. Paper dan artikel tentang berbagai masalah mental health, hal ini termasuk: skizofrenia, depresi, kecemasan, gangguan bipolar, dan PTSD. Bagian ini telah dibagi dan dikategorikan. Keefektifan proses pembelajaran mesin yang digunakan juga telah diperhatikan karena mungkin bermanfaat bagi industri medis dalam penambangan data atau domain data besar. Akurasi adalah proporsi jumlah prediksi yang benar dibandingkan dengan total jumlah prediksi. F-measure adalah harmonic mean antara precision dan recall, yang menggabungkan informasi tentang false positives, false negatives, true positives, dan true negatives. Sedangkan AUC (Area Under the Curve) adalah area di bawah kurva ROC (Receiver Operating Characteristic), yang memplotkan true positive



rate (y-axis) versus false positive rate (x-axis). Berdasarkan Tabel 2.1, dapat dilihat bahwa beberapa algoritma memiliki tingkat akurasi yang cukup tinggi seperti Random subspace (87.5%), J48 (87.8%), Random forest (89.0%), dan Random tree (85.1%). Namun, AUC juga perlu dipertimbangkan karena menunjukkan kemampuan algoritma untuk membedakan kelas positif dan negatif. Algoritma dengan AUC tertinggi adalah Random forest (94.3%), diikuti oleh Bayesian network (88.9%), Random subspace (91.7%), K-star (81.4%), dan Naive Bayes (85.3%). Sedangkan Logistic regression dan Sequential minimal optimisation memiliki performa terendah dalam hal semua metrik evaluasi.

Penelitian terdahulu selanjutnya dengan algoritma Whale Optimization Algorithm (WOA) yang dipadukan dengan K-Nearest Neighbors (KNN) dilakukan oleh Hosseinabadi et al. (2022). Case study berupa medical diagnostic bagian prediksi dengan tingkat akurasi mencapai 91.29%. Tujuan dari penelitian ini adalah untuk meningkatkan kinerja diagnostik medis dengan menggabungkan klasifikasi yang membutuhkan pengawasan dan varian diskrit dari teknik optimasi meta-heuristik WOA. Peneliti mendesain WOA diskrit dengan mendefinisikan kembali komponen yang relevan untuk area diskrit dan membuat fungsi eksplorasi baru dengan pilihan yang lebih luas. Berdasarkan akurasi klasifikasi pengklasifikasi KNN, peran WOA dimasukkan. Menurut hasil temuan pada eksperimental, teknik WOA+KNN meningkatkan efisiensi algoritma machine learning. Penelitian terbaru yang dilakukan oleh Hosseinabadi et al. (2022) secara mendetail memaparkan tentang pengembangan teknik diagnostik medis menggunakan algoritma Whale Optimization Algorithm (WOA) dan K-Nearest Neighbors (KNN). Penelitian ini

bertujuan untuk meningkatkan kinerja diagnostik medis dengan menggabungkan teknik klasifikasi yang membutuhkan pengawasan dan varian diskrit dari teknik optimasi meta-heuristik WOA. Tingkat akurasi yang dicapai dalam penelitian ini mencapai 91.29%. Peneliti mendesain WOA diskrit dengan mendefinisikan kembali komponen yang relevan untuk area diskrit dan membuat fungsi eksplorasi baru dengan pilihan yang lebih luas. Dalam penelitian ini, peran WOA dimasukkan dalam klasifikasi pengklasifikasi KNN. Berdasarkan hasil eksperimental, teknik WOA+KNN meningkatkan efisiensi algoritma machine learning dalam diagnostik medis.

Kemudian penelitian yang dilakukan oleh Mat et al. (2021) dengan WOA yang dikombinasikan dengan Levy Flight (LF), yang didasarkan pada metode perburuan paus, dan metode pencarian Levy flight (LF) digunakan untuk menerapkan pengelompokan. Pencarian global algoritma WOA telah ditingkatkan, dan hasil pengelompokan telah diperbaiki, menggunakan pendekatan penerbangan Levy. Keefektifan WOALF telah dibandingkan dengan WOA asli dan tiga algoritma pengelompokan fundamental (k-means, k-medoids, dan fuzzy cmeans). Dibandingkan dengan algoritma ini, telah ditunjukkan bahwa WOA-LF memberikan hasil akhir yang lebih baik secara keseluruhan. Karena LF kinerja WOA dalam masalah pengelompokan, ini dapat digunakan sebagai algoritma pengelompokan pengganti.

Penelitian oleh Tawhid & Ibrahim (2019) menggunakan algoritma kombinasi Binary Whale Optimization Algorithm (BWOA) berfokus ke feature selection (FS) pada berbagai macam jenis dataset. Dengan perbandingan data

training 50 dan data validation 50, BWOA menghasilkan akurasi rata-rata sebesar 78%, 72%, dan 83% dari 18 dataset yang digunakan. Efektivitas pengenalan dan klasifikasi dapat sangat dipengaruhi oleh pemilihan fitur, yang merupakan bagian penting dalam penelitian ini. Selain itu, dapat digunakan juga pada algoritma natural seperti PSO, DE, ataupun GA. BWOA digunakan untuk memilih fitur-fitur yang paling signifikan dalam dataset yang digunakan untuk meningkatkan efektivitas pengenalan dan klasifikasi. Dalam penelitian ini, perbandingan data training 50 dan data validasi 50 digunakan untuk menguji akurasi rata-rata BWOA pada 18 dataset yang berbeda. Selain itu, penelitian ini juga menunjukkan bahwa pemilihan fitur merupakan bagian penting dalam meningkatkan efektivitas pengenalan dan klasifikasi pada dataset. Algoritma BWOA juga dapat digunakan pada algoritma natural seperti PSO, DE, ataupun GA untuk meningkatkan efisiensi dan akurasi dalam pemilihan fitur.

Bidang lain yang menerapkan kombinasi WOA adalah penelitian yang dilakukan oleh Sumathi & Malarvizhi (2022). Whale Optimization Algorithm with Long Short-Term Memory (WOA-LSTM). Survei dilakukan dengan menggunakan serangkaian kuesioner untuk manajemen kualitas perangkat lunak dan prosedur penjaminan. Dalam hal ini, proses rekomendasi menggunakan model LSTM. Performa meningkat sebagai hasil dari penyetelan hyperparameter model LSTM yang sangat baik dari WOA. Dalam sampel ekstrem terkait, model WOA-LSTM menghasilkan penilaian khusus konteks dan saran umum untuk rekomendasi aplikasi. Untuk memastikan kinerja efektif model WOA-LSTM, sejumlah pengujian dijalankan, dan hasilnya menyoroti hasil efektif untuk saran penulis.

Digunakannya kombinasi WOA dan LSTM untuk meningkatkan kinerja model dalam proses rekomendasi manajemen kualitas perangkat lunak dan prosedur penjaminan. Dalam penelitian ini, WOA digunakan untuk menyetel hyperparameter model LSTM, sehingga menghasilkan model yang lebih baik dalam melakukan rekomendasi aplikasi. Hasil pengujian menunjukkan bahwa model WOA-LSTM menghasilkan penilaian khusus konteks dan saran umum yang efektif dalam proses rekomendasi aplikasi. Penelitian ini menunjukkan potensi penggunaan kombinasi WOA dan LSTM dalam bidang manajemen kualitas perangkat lunak dan prosedur penjaminan.

Berdasarkan dari keseluruhan penelitian yang telah dibahas dalam sesi ini, dapat disimpulkan bahwa WOA termasuk kedalam natural algorithm yang cocok digunakan terutama prediksi di dunia medic untuk dikombinasikan dengan algoritma KNN menjadi KNN-WOA. Implementasi KNN-WOA di prediksi mental health belum ditemukan. Pengembangan prosedur pada KNN-WOA diperlukan karena dirasa dapat meningkatkan akurasi dan tentunya kepercayaan atau trust dapat ikut meningkat juga.

## 2.2. Keaslian Penelitian

Tabel 2.2 Matriks literatur review dan posisi penelitian  
Diagnosa Kesehatan Mental Menggunakan K-Nearest Neighbor dengan Whale Optimization Algorithm

| No | Judul   | Peneliti, Media Publikasi, dan Tahun              | Tujuan Penelitian   | Kesimpulan  | Saran atau Kekurangan   | Perbandingan   |
|----|---|---|---|---|---|--|
| 1  | Whale Optimization for Wavelet-Based Unsupervised Medical Image Segmentation: Application to CT and MR Images | Vaiyapuri, T. & Alaskar, H., 2020, Atlantis Press | Karya ini bertujuan untuk mengatasi masalah ini mengintegrasikan kekuatan multiresolusi analisis dan teknik optimisasi meta-heuristik untuk segmentasi citra medis tanpa pengawasan | Penelitian ini berkontribusi untuk mengisi kesenjangan dalam literatur menyelidiki untuk pertama kalinya algoritma kecerdasan yang berbeda seperti fuzzy, algoritma genetika (GA) dan partikel swarm optimasi (PSO) untuk segmentasi citra medis tanpa pengawasan untuk menunjukkan kemanjuran pendekatan yang diusulkan. | Bangun untuk akhiran proses WOA, ketika salah satu dari kriteria berikut terpenuhi: (i) solusi optimal global ditemukan, (ii) maksimum iterasi tercapai, (iii) kebugaran semua paus sama. | Penelitian Vaiyapuri mengimplementasikan WOA pada CT dan MR Images, penelitian kali ini mengimplementasikan WOA pada mental health |
| 2  | Enhanced whale optimization   | Shahraki, M. H. N.,                               | Mengusulkan sebuah  | Algoritme metaheuristik   | Mengembangkan algoritma yang  | Implementasi WOA pada deteksi Covid-19 pada  |

Tabel 2.2 (Lanjutan) Matriks literatur review dan posisi penelitian  
Diagnosa Kesehatan Mental Menggunakan K-Nearest Neighbor dengan Whale Optimization Algorithm

| No | Judul   | Peneliti, Media Publikasi, dan Tahun  | Tujuan Penelitian  | Kesimpulan   | Saran atau Kekurangan   | Perbandingan  |
|----|---|---|--|--|---|---|
|    | algorithm for medical feature selection: A COVID-19 case study  | Zamani, H., & Mirjalili, S., 2022, Elsevier: Computers in Biology and Medicine                      | algoritme pengoptimalan paus yang ditingkatkan bernama E-WOA menggunakan mekanisme penyatuan dan tiga pencarian efektif strategi bernama migrasi, pemilihan preferensial, dan memperkaya mangsa yang mengelilingi. | menunjukkan bahwa mereka efisien untuk mendekati subset fitur yang optimal dalam waktu yang wajar.   | diusulkan untuk pemecahan masalah pemilihan fitur tanpa pengawasan dunia nyata menggunakan operator heuristik.  | penelitian ini. Penulis akan memanfaatkan algoritma ini kedalam mental health detection.  |
| 3  | Research on multi-modal autonomous diagnosis algorithm of COVID-19 based on whale optimized support vector machine and improved D-S evidence fusion | Wang, G., Gio, S., Han, L., Song, X., & Zhao, Y., 2022, Elsevier: Computers in Biology and Medicine | Metode pengenalan fitur sinyal gambar CT tunggal dalam diagnosis mandiri penyakit tidak dapat secara akurat dan andal mengklasifikasikan COVID-19, dan mudah dikacaukan  | Metode mengganti parameter fitur tunggal dengan indikator multi-modal (CT, indeks laboratorium rutin, sitokin serum dan kemokin) memberikan gambaran lebih lanjut sumber sinyal yang andal untuk model | Sebaiknya penelitian lebih berfokus pada pencocokan optimal dari berbagai indikator diagnostik dengan biaya rendah untuk meningkatkan efek aplikasi teori fusi D-S di pengakuan COVID-19. | Implementasi dari WOA yang di kembangkan pada diagnostic Covid-19. Peneliti memfokuskan implementasi WOA pada diagnostic mental health. |

Tabel 2.2 (Lanjutan) Matriks literatur review dan posisi penelitian  
 Diagnosa Kesehatan Mental Menggunakan K-Nearest Neighbor dengan Whale Optimization Algorithm

| No | Judul   | Peneliti, Media Publikasi, dan Tahun  | Tujuan Penelitian  | Kesimpulan   | Saran atau Kekurangan  | Perbandingan   |
|----|---|---|--|--|--|--|
|    |   |   | dengan kasus yang dicurigai.   | diagnosis, yang secara efektif dapat membedakan COVID-19 dan yang dicurigai kasus.   |  |  |
| 4  | Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges | Chung, J. & Teo, J., 2022, Hindawi: Applied Computational Intelligence and Soft Computing | Studi ini bertujuan untuk mengumpulkan data dan meningkatkan analisis dari fitur teks dan suara. | Banyak teknik dan algoritma yang berbeda telah diperkenalkan dan mengusulkan untuk menguji dan memecahkan kesehatan mental) Sebagai mengklasifikasikan data kesehatan mental umumnya merupakan masalah yang sangat menantang, fitur yang digunakan dalam algoritma pembelajaran mesin akan secara signifikan | Masih banyak solusi yang bisa disempurnakan. Selain itu, masih banyak masalah yang harus ditemukan dan diuji menggunakan berbagai macam pengaturan di mesin belajar untuk domain kesehatan mental. | Prediksi mental health dilakukan dengan berbagai algoritma, belum digunakan WOA dalam implementasinya. Peneliti kali ini akan menggunakan WOA dalam memprediksi mental health pada pasien. |

Tabel 2.2 (Lanjutan) Matriks literatur review dan posisi penelitian  
 Diagnosa Kesehatan Mental Menggunakan K-Nearest Neighbor dengan Whale Optimization Algorithm

| No | Judul   | Peneliti, Media Publikasi, dan Tahun  | Tujuan Penelitian  | Kesimpulan  | Saran atau Kekurangan  | Perbandingan   |
|----|---|---|--|---|--|--|
|    |   |   |  | mempengaruhi kinerja klasifikasi.   |  |  |
| 5  | Whale Optimization-based Prediction for Medical Diagnostic          | Hossinabadi, A. A. R., Sadeghilalimi, M., Sharch, M. B., Mouhoub, M., & Sadaoui, S., 2022, Proceedings of the 14th International Conference on Agents and Artificial Intelligence (ICAART 2022) | Meningkatkan akurasi pendeteksian penyakit dengan menggabungkan versi terpisah dari Whale Optimization Algoritma (WOA) ke dalam kerangka kerja klasifikasi terawasi (KNN). | Eksperimental hasil menunjukkan bahwa pendekatan WOA+KNN meningkatkan kinerja pembelajaran mesin algoritma. | Salah satu arah penelitian yang menarik namun menantang adalah untuk menggabungkan WOA dengan teknik terinspirasi alam lainnya | WOA di implementasikan untuk prediksi medis. Peneliti memfokuskan tidak hanya medis secara umum namun medis terkhusus mental health. |
| 6  | An application of the whale optimization algorithm with Levy flight | Mat, A. N., Inan, O., & Karakoyun, M., 2021, An International Journal of Optimization   | Untuk mencapai hasil yang lebih baik dengan solusi sederhana dan mengelompokkan  | Membandingkan algoritma pengelompokan WOA-LF dengan hybrid pengelompokan pendekatan atau                    | Di sisi lain, algoritma dapat ditingkatkan dan diterapkan pada optimasi yang berbeda masalah.                                  | WOA dan FL-WOA di implementasikan untuk mengklasifikasikan berbagai dataset kesehatan, peneliti fokus kepada mental health.          |



Tabel 2.2 (Lanjutan) Matriks literatur review dan posisi penelitian  
 Diagnosa Kesehatan Mental Menggunakan K-Nearest Neighbor dengan Whale Optimization Algorithm

| No | Judul  | Peneliti, Media Publikasi, dan Tahun  | Tujuan Penelitian   | Kesimpulan   | Saran atau Kekurangan  | Perbandingan   |
|----|--|---|---|--|--|--|
|    | strategy for clustering of medical datasets  | and Control: Theories & Applications  | data yang tidak berlabel.   | menggunakan fungsi kebugaran lainnya selama pengelompokan mungkin menjadi subjek masa depan riset.   |  |  |
| 7  | Feature selection based on rough set approach, wrapper approach, and binary whale optimization algorithm | Tahwid, M. A. & Ibrahim, A. M., 2019. International Journal of Machine Learning and Cybernetics | Untuk bidang penelitian yang menjanjikan sekitar diperoleh solusi terbaik sehingga kita dapat menemukan solusi optimal / reduksi minimum. | Algoritma yang tulisan ini usulkan dapat diterapkan untuk pemilihan subset fitur gambar. Juga, BWOA juga bisa diperluas untuk menghibridisasi algoritme kecerdasan segerombolan canggih seperti PSO, DE, GA dan pengoptimalan pencarian kukuk. | Merancang algoritma feature selection (FS) bergantung pada evaluasi pengukuran dan strategi pencarian. | Pengembangan BWOA yang berfokus pada feature selection. Sedangkan penulis memfokuskan diri pada implementasi WOA pada mental health. |
| 8  | Whale Optimization Algorithm with Deep Learning-   | Sumathi, P. & Malarvizhi, N., 2022, International   | Dapat mengembangkan Algoritma Pengoptimalan   | Survei dilakukan dengan menggunakan seperangkat kuesioner untuk  | Model WOA-LSTM menghasilkan penilaian khusus konteks dan saran   | WOA-LSTM langsung di implementasikan berupa medical mobile. Peneliti kali ini memfokuskan  |

Tabel 2.2 (Lanjutan) Matriks literatur review dan posisi penelitian  
 Diagnosa Kesehatan Mental Menggunakan K-Nearest Neighbor dengan Whale Optimization Algorithm

| No | Judul   | Peneliti, Media Publikasi, dan Tahun         | Tujuan Penelitian  | Kesimpulan  | Saran atau Kekmahan   | Perbandingan                    |
|----|---|--|--|---|---|---------------------------------|
|    | Based Usability Recommendation Model for Medical Mobile | Journal of Engineering Trends and Technology | Paus baru dengan memori jangka pendek panjang (WOA-LSTM) Model Rekomendasi Kegunaan Berbasis untuk Aplikasi Seluler Medis. | jaminan kualitas perangkat lunak dan praktik manajemen kualitas perangkat lunak. Dalam hal ini, model LSTM digunakan untuk proses rekomendasi. Hyperparameter yang terlibat dalam model LSTM disetel secara optimal oleh WOA, menghasilkan peningkatan kinerja. | umum untuk rekomendasi Aplikasi dalam sampel ekstrem terkait. | implementasi di belakang layar. |

### 2.3. Landasan Teori

#### 2.3.1 Mental Health

Kesehatan mental atau mental health dalam bahasa Indonesia dapat didefinisikan sebagai kondisi kesehatan yang meliputi kesehatan psikologis dan emosional seseorang (Dewi, 2012). Hal ini mencakup berbagai aspek kehidupan seseorang, seperti pemikiran, perasaan, perilaku, dan hubungan. Kesehatan mental yang baik berarti memiliki kemampuan untuk mengelola stres kehidupan sehari-hari, mengatasi tantangan, dan menikmati pengalaman positif. Ini juga melibatkan ketiadaan gangguan atau penyakit mental, seperti depresi, kecemasan, gangguan bipolar, dan skizofrenia (Menteri Kesehatan Republik Indonesia, 2012). Kesehatan mental merupakan komponen penting dari kesehatan dan kesejahteraan secara keseluruhan, dan penting untuk mengambil langkah-langkah untuk menjaga kesehatan mental yang baik sepanjang hidup seseorang.

Mendeteksi kesehatan mental seseorang memiliki banyak manfaat dan penting untuk dilakukan karena:

- (1) Mencegah masalah kesehatan mental yang lebih serius: Dengan mendeteksi masalah kesehatan mental seseorang secara dini, dapat mencegah kondisi tersebut memburuk menjadi lebih serius atau bahkan menyebabkan gangguan mental yang lebih berat. Deteksi dini masalah kesehatan mental sangat penting untuk mencegah kondisi tersebut semakin parah atau bahkan berkembang menjadi gangguan mental yang lebih serius. Semakin cepat seseorang mendapatkan diagnosis dan pengobatan yang tepat, semakin baik juga prospek pemulihan dari kondisi kesehatan mental yang dialaminya. Oleh karena itu,

sangat penting untuk meningkatkan kesadaran dan edukasi mengenai kesehatan mental, serta mengembangkan alat dan teknologi untuk mendeteksi masalah kesehatan mental sejak dini.

- (2) Meningkatkan kualitas hidup: Kondisi kesehatan mental yang baik dapat meningkatkan kualitas hidup seseorang dengan mengurangi tingkat stres, kecemasan, dan depresi yang dapat mempengaruhi kemampuan seseorang untuk menikmati hidup. Orang dengan kesehatan mental yang baik dapat merasa lebih bahagia, percaya diri, dan dapat mengatasi tantangan hidup dengan lebih baik. Di sisi lain, stres, kecemasan, dan depresi yang berkelanjutan dapat menyebabkan masalah kesehatan mental yang lebih serius, bahkan mempengaruhi kesehatan fisik seseorang. Oleh karena itu, penting untuk merawat kesehatan mental dengan cara yang sama seperti merawat kesehatan fisik.
- (3) Meningkatkan produktivitas: Orang yang memiliki kesehatan mental yang baik dapat menjadi lebih produktif dan efisien dalam pekerjaan dan kehidupan sehari-hari, karena mereka dapat mengatasi tantangan dengan lebih baik dan memiliki kemampuan untuk fokus dan berkonsentrasi yang lebih baik. Kesehatan mental yang baik memainkan peran penting dalam produktivitas dan efisiensi seseorang di berbagai aspek kehidupan. Dalam lingkungan kerja, orang yang memiliki kesehatan mental yang baik cenderung lebih produktif dan mampu menyelesaikan tugas-tugas mereka dengan lebih efisien. Mereka memiliki kemampuan untuk menyelesaikan masalah dengan baik, mengambil keputusan yang tepat, dan bekerja dalam tim dengan baik. Di luar pekerjaan,

kesehatan mental yang baik juga dapat membantu seseorang dalam melakukan tugas-tugas sehari-hari dengan lebih mudah dan efisien. Mereka dapat mengatasi tantangan kehidupan dengan lebih baik dan memiliki kemampuan untuk fokus dan berkonsentrasi yang lebih baik. Selain itu, kesehatan mental yang baik juga dapat meningkatkan hubungan sosial dan interaksi dengan orang lain, yang dapat membantu seseorang merasa lebih terhubung dan berada dalam lingkungan yang sehat secara mental.

- (4) Mengurangi stigmatisasi: Mendeteksi masalah kesehatan mental seseorang secara dini dan memberikan dukungan yang tepat dapat membantu mengurangi stigmatisasi dan diskriminasi terhadap orang yang memiliki masalah kesehatan mental. Hal ini karena dengan mendeteksi masalah kesehatan mental pada tahap awal, seseorang dapat menerima pengobatan atau terapi yang tepat sebelum kondisi menjadi lebih serius dan memerlukan penanganan yang lebih ekstensif. Selain itu, dukungan yang tepat dari keluarga, teman, atau profesional kesehatan mental juga dapat membantu seseorang dalam mengatasi masalah kesehatan mental dan memperbaiki kualitas hidupnya. Dengan demikian, stigmatisasi dan diskriminasi terhadap orang dengan masalah kesehatan mental dapat dikurangi karena mereka menerima dukungan dan pengobatan yang tepat.
- (5) Meningkatkan kesadaran dan pemahaman tentang kesehatan mental: Dengan meningkatkan kesadaran dan pemahaman tentang kesehatan mental, masyarakat dapat lebih memahami pentingnya menjaga kesehatan mental yang baik dan mencari bantuan saat diperlukan. Hal tersebut sangat penting karena

masih banyak masyarakat yang kurang memahami tentang pentingnya menjaga kesehatan mental dan masih ada stigma negatif terhadap orang dengan masalah kesehatan mental. Dengan meningkatkan kesadaran dan pemahaman tentang hal ini, diharapkan masyarakat dapat lebih terbuka dan mendukung orang-orang yang membutuhkan bantuan dalam menjaga kesehatan mental mereka. Selain itu, hal ini juga dapat mendorong pemerintah dan lembaga kesehatan untuk memberikan perhatian yang lebih besar terhadap masalah kesehatan mental dan meningkatkan pelayanan kesehatan mental bagi masyarakat.

Dengan demikian, deteksi dini masalah kesehatan mental sangat penting untuk menjaga kesehatan dan kesejahteraan secara keseluruhan. Deteksi dini masalah kesehatan mental dapat membantu seseorang untuk mendapatkan perawatan dan dukungan yang tepat pada tahap awal sebelum kondisi tersebut memburuk dan berdampak negatif pada kesehatan secara keseluruhan. Hal ini juga dapat membantu mencegah terjadinya komplikasi kesehatan fisik yang terkait dengan kondisi kesehatan mental, serta meningkatkan kualitas hidup secara keseluruhan.

### **2.3.2 WOA**

Whale Optimization Algorithm (WOA) ialah Algoritma berbasis populasi yang disebut WOA menggunakan perburuan kolektif paus bungkuk untuk melambangkan solusi alternatif (Mirjalili dan Lewis, 2016; Sangaiah et al., 2020). Dengan cara ini, paus membuat gelembung berbentuk spiral sebagai bagian dari kebiasaan mencari makan. Paus melakukan perjalanan dengan cara

eksploitasi/eksplorasi menuju mangsanya (Mirjalili dan Lewis, 2016). Dengan kata lain Whale Optimization Algorithm (WOA) adalah algoritma yang berbasis populasi dan terinspirasi dari perilaku paus bungkuk dalam mencari makan. Algoritma ini bekerja dengan cara memperbarui posisi setiap individu dalam populasi berdasarkan gerakan spiral, menurun, dan melingkar untuk melakukan eksplorasi dan eksploitasi terhadap solusi alternatif.

Gerakan spiral, menurun, dan melingkar digunakan untuk mengeksploitasi. Yang pertama, setiap paus melingkari targetnya saat mendekati. Dalam teknik kedua, setiap paus mengikuti kurva spiral saat bergerak menuju sarangnya. WOA menghasilkan solusi yang optimal dengan cara mengubah parameter dalam ruang pencarian. Hal ini dilakukan dengan cara mengevaluasi kinerja setiap individu dalam populasi dan memperbarui posisi mereka berdasarkan persamaan yang didefinisikan dengan asumsi bahwa posisi mangsa sama. Setiap paus akan memperbarui lokasinya selama gerakan mengecil mengecil menggunakan persamaan berikut yang didefinisikan dengan asumsi bahwa posisi mangsanya sama atau mirip dengan Best Whale (BW) yang ditunjukkan pada Persamaan 1 dan Persamaan 2.

WOA telah digunakan dalam berbagai aplikasi seperti pengoptimalan parameter, prediksi, klasifikasi, dan feature selection pada berbagai jenis dataset. Algoritma ini telah terbukti efektif dan dapat bersaing dengan algoritma optimasi lainnya seperti Particle Swarm Optimization (PSO), Differential Evolution (DE), dan Genetic Algorithm (GA) (Mirjalili dan Lewis, 2016).

$$D = |CX^*(t) - X(t)| \dots \dots \dots (1)$$

$$X(t + 1) = X^*(t) - AD \dots \dots \dots (2)$$

Iterasi saat ini dan selanjutnya dalam contoh di atas masing-masing dilambangkan dengan  $t$  dan  $t+1$ . Posisi BW dan paus tertentu masing-masing diwakili oleh  $X$  dan  $X_i$ . Ini adalah perhitungan untuk A dan C.

$$A = 2ar - a \dots\dots\dots(3)$$

$$C = 2r \dots\dots\dots(4)$$

Setelah mendapatkan posisi A dan C pada iterasi  $t+1$ , dilakukan perhitungan nilai fungsi objektif pada masing-masing posisi tersebut. Jika nilai fungsi objektif pada posisi A lebih baik daripada posisi  $X_i$ , maka posisi  $X_i$  diperbarui dengan posisi A. Begitu pula jika nilai fungsi objektif pada posisi C lebih baik daripada posisi BW, maka posisi BW diperbarui dengan posisi C. Setelah itu, dilakukan iterasi berikutnya dengan menggunakan posisi BW dan paus-paus lainnya yang telah diperbarui tersebut.

Dalam rentang  $[0,2]$  dan  $[0,1]$ , masing-masing,  $a$  dan  $r$  adalah parameter acak. Untuk mencapai penyusutan, amati penurunan dari 2 menjadi 0 pada setiap siklus. Persamaan berikut (Mirjalili dan Lewis, 2016) menghasilkan gerak spiral.

$$X(t+1) = D'e^{bt} \cos(2\pi l) + X^*(t) \dots\dots\dots(5)$$

Pada contoh di atas,  $b$  adalah konstanta yang digunakan untuk mengatur kurva spiral, dan  $l$  adalah nilai acak dalam rentang  $[-1,1]$ . Rumus berikut digunakan untuk menghitung jarak  $D'$  antara setiap paus dan mangsanya: (Mirjalili dan Lewis, 2016). Dalam rumus ini, jarak antara paus dan mangsa dihitung sebagai produk antara jarak Euclidean dan perbedaan antara konstanta  $b$  dan nilai acak  $l$ . Nilai  $b$  digunakan untuk mengatur ketat atau longgar kurva spiral, sedangkan nilai  $l$  digunakan untuk menentukan jarak paus dari pusat kurva spiral. Dengan mengubah



nilai  $b$  dan  $l$ , kita dapat menghasilkan kurva spiral dengan berbagai bentuk dan ukuran yang berbeda, sehingga memungkinkan untuk mengoptimalkan strategi perburuan paus dalam menangkap mangsa.

$$D' = |X^*(t) - X(t)| \dots \dots \dots (6)$$

Alih-alih menuju BW seperti yang terjadi pada gerakan spiral dan lingkungan menyusut yang dijelaskan di atas, setiap paus mencari mangsa secara acak dengan memperbarui posisinya sesuai dengan paus yang dipilih secara acak. Berikut ini adalah persamaan yang relevan untuk gerakan acak (Mirjalili dan Lewis, 2016). Strategi perburuan acak ini berbeda dengan strategi gerakan spiral yang dijelaskan sebelumnya, di mana setiap paus mengikuti kurva spiral yang ditentukan oleh nilai konstanta  $b$  dan nilai acak  $l$ . Dalam strategi gerakan acak, setiap paus mencari mangsa dengan mengambil langkah acak pada setiap iterasi, yang menghasilkan gerakan yang kurang terduga dan dapat membantu paus menemukan mangsa dengan lebih efektif.

$$D = |CX_{rand}(t) - X(t)| \dots \dots \dots (7)$$

$$X(t + 1) = X_{rand}(t) - AD \dots \dots \dots (8)$$

$X_{rand}$  adalah paus yang dipilih secara acak dalam skenario di atas. Algoritma WOA beroperasi sebagai berikut mengingat persamaan yang disebutkan di atas yang mengarahkan taktik eksploitasi dan eksplorasi.  $P$  dipilih secara acak dari rentang  $[0, 1]$ . Jika  $p$  lebih besar dari atau sama dengan  $0,5$ , paus akan berputar. Jika tidak, kita harus menentukan nilai  $A$ . Ini akan melakukan operasi pengecilan penyusutan jika  $|A|$  kurang dari atau sama dengan  $1$ . Jika tidak, paus akan bergoyang secara acak seperti yang dikatakan sebelumnya. Dalam algoritma WOA

(Whale Optimization Algorithm), paus yang dipilih secara acak diwakili oleh variabel  $X_{rand}$ , yang merupakan salah satu dari tiga paus dalam populasi yang sedang dioptimalkan.  $X_{rand}$  digunakan untuk mengarahkan taktik eksploitasi dan eksplorasi yang digunakan dalam algoritma ini. Setiap iterasi dalam algoritma WOA terdiri dari dua tahap: tahap eksplorasi dan tahap eksploitasi. Tahap eksplorasi digunakan untuk mencari wilayah pencarian yang lebih luas dalam ruang pencarian solusi, sedangkan tahap eksploitasi digunakan untuk mencari solusi optimal dalam wilayah pencarian yang lebih sempit.

Perlu diingat bahwa nilai  $A$  mungkin lebih besar dari 1 pada awal proses, memungkinkan eksplorasi tambahan. Oleh karena itu, nilai  $A$  akan sering menjadi kurang dari 1, yang akan mendorong penyusutan lebih lanjut seiring dengan penurunan nilai. Ini mewakili perilaku algoritme, yang dimulai dengan lebih banyak eksplorasi dan diakhiri dengan lebih banyak eksploitasi. Mengingat bahwa WOA pertama kali dikembangkan untuk mengatasi masalah pengoptimalan berkelanjutan (Mirjalili dan Lewis, 2016), kami harus mengadaptasinya ke ruang diskrit untuk menyesuaikan bobot bobot fitur dataset pasien. Bagian selanjutnya memberikan deskripsi tentang bagaimana berbagai komponen WOA didiskritisasi.

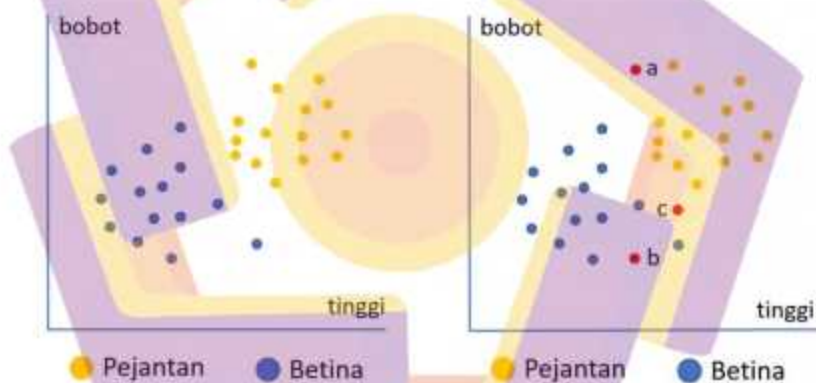
### **2.3.3 Algoritma Pembentuk Deteksi Mental Health**

#### **2.3.3.1 KNN**

KNN (K-Nearest Neighbors) adalah salah satu algoritma machine learning yang digunakan untuk klasifikasi dan regresi. Algoritma ini bekerja dengan cara mencari  $k$  titik data terdekat dari data yang ingin diklasifikasikan (dalam kasus

klasifikasi) atau data yang ingin diprediksi (dalam kasus regresi). Kemudian, kelas atau nilai regresi dari data yang ingin diklasifikasikan atau diprediksi dihitung dengan cara mengambil mayoritas kelas atau rata-rata nilai regresi dari  $k$  titik data terdekat.

Algoritma KNN (K-Nearest Neighbors) adalah salah satu algoritma klasifikasi yang populer digunakan dalam machine learning. Algoritma ini bekerja dengan mencari  $k$  titik data terdekat dengan titik data yang sedang diprediksi, dan kemudian mengklasifikasikan titik data tersebut berdasarkan mayoritas label  $k$  titik data terdekat seperti pada Gambar 2.1.



Gambar 2.1 Contoh Kasus KNN

Gambar 2.1 yang diambil dari website [sainsdata.id](http://sainsdata.id) memberikan data training disebelah kiri dan data testing disebelah kanan dengan bintik warna merah. Misalkan bintik merah a akan dicari tahu merupakan pejantan (bintik kuning) atau betina (bintik biru). Ketika ditarik garis untuk menghitung jarak antara bintik a dengan bintik tetangga terdekatnya, dengan memperhatikan bobot serta jarak maka bintik merah a dapat diklasifikasikan sebagai pejantan karena alasan tersebut.

Algoritma KNN dapat digunakan pada data numerik dan kategorikal. Pada data numerik, jarak antara titik data dapat dihitung menggunakan metrik jarak, seperti Euclidean distance atau Manhattan distance. Pada data kategorikal, jarak antara titik data dapat dihitung menggunakan metrik jarak Hamming.

Algoritma KNN dapat digunakan pada data numerik dan kategorikal. Pada data numerik, jarak antara titik data dapat dihitung menggunakan metrik jarak, seperti Euclidean distance atau Manhattan distance. Euclidean distance mengukur jarak antara dua titik dalam bentuk garis lurus, sedangkan Manhattan distance mengukur jarak antara dua titik dalam bentuk garis tegak lurus. Perbedaan antara kedua metrik jarak tersebut terletak pada cara menghitung jarak antara atribut numerik pada titik data.

Pada data kategorikal, jarak antara titik data dapat dihitung menggunakan metrik jarak Hamming. Metrik jarak Hamming mengukur jarak antara dua titik data dengan cara menghitung jumlah atribut kategorikal yang berbeda pada kedua titik data. Metrik jarak Hamming hanya berlaku pada data kategorikal biner, di mana dalam kasus ini, nilai atribut hanya dapat bernilai 0 atau 1.

Selain itu, pada data yang memiliki kombinasi atribut numerik dan kategorikal, dapat digunakan metrik jarak gabungan seperti Gower distance atau Minkowski distance. Metrik jarak gabungan ini menggabungkan metrik jarak numerik dan metrik jarak kategorikal untuk menghitung jarak antara titik data.

Dalam penggunaan algoritma KNN, pemilihan metrik jarak yang tepat sangat penting untuk menghasilkan model klasifikasi yang akurat. Pemilihan metrik

jarak yang tepat harus disesuaikan dengan tipe data yang digunakan dan karakteristik dari setiap atribut pada data tersebut (Hajihashemi et al., 2019).

Parameter  $k$  pada KNN menentukan jumlah titik data terdekat yang akan diambil untuk menghitung kelas atau nilai regresi. KNN biasanya membutuhkan preprocessing data untuk menghilangkan nilai-nilai yang hilang atau data yang tidak relevan serta standarisasi nilai numerik agar tidak mempengaruhi hasil klasifikasi atau regresi. Algoritma KNN adalah algoritma yang mudah dimengerti dan diimplementasikan. Namun, kelemahan dari algoritma ini adalah kemampuan komputasi yang tinggi pada data yang besar dan sensitivitas pada data yang memiliki banyak fitur atau dimensi.

Parameter  $k$  pada KNN (K-Nearest Neighbors) menentukan jumlah titik data terdekat yang akan diambil untuk menghitung kelas atau nilai regresi pada titik data yang sedang diprediksi. Semakin besar nilai  $k$ , semakin banyak titik data terdekat yang diambil untuk melakukan klasifikasi atau regresi, namun semakin kecil nilai  $k$ , semakin sensitif model terhadap noise atau variabilitas pada data.

KNN biasanya membutuhkan preprocessing data untuk menghilangkan nilai-nilai yang hilang atau data yang tidak relevan serta standarisasi nilai numerik agar tidak mempengaruhi hasil klasifikasi atau regresi. Preprocessing data seperti ini diperlukan untuk memastikan bahwa hasil klasifikasi atau regresi yang dihasilkan akurat dan terpercaya.

Preprocessing data dapat dilakukan dengan berbagai cara, misalnya dengan menghilangkan baris data yang tidak memiliki nilai atau menggantinya dengan nilai rata-rata atau nilai yang paling sering muncul pada kolom tersebut. Selain itu, untuk

data numerik, nilai-nilai dapat distandarisasi atau dinormalisasi agar memiliki rentang nilai yang serupa. Dalam hal ini, nilai-nilai pada setiap atribut numerik diubah sedemikian rupa sehingga memiliki nilai rata-rata nol dan variansi satu. Dalam beberapa kasus, pemilihan nilai  $k$  yang tepat dapat mempengaruhi hasil klasifikasi atau regresi yang dihasilkan oleh model KNN. Untuk memilih nilai  $k$  yang optimal, dapat dilakukan teknik *cross-validation* dengan membagi data menjadi beberapa bagian dan melakukan validasi pada setiap bagian tersebut. Selain itu, dapat juga menggunakan teknik *Grid Search* untuk mencari nilai  $k$  yang optimal dengan mengujinya pada kisaran nilai yang telah ditentukan (Hajihashemi et al., 2019).

KNN dapat digunakan sebagai salah satu metode untuk mendeteksi tingkat kesehatan mental seseorang. Algoritma ini dapat digunakan untuk memprediksi kondisi kesehatan mental berdasarkan data klinis dan non-klinis, seperti gejala yang dialami, riwayat medis, riwayat keluarga, dan faktor sosial.

Untuk menggunakan KNN dalam mendeteksi kesehatan mental, data dapat dipreproses dan diekstraksi dari sumber yang relevan, seperti rekam medis, survei kesehatan, atau kuesioner online. Data ini kemudian dapat dianalisis menggunakan algoritma KNN dengan memasukkan variabel atau fitur yang relevan, seperti gejala, diagnosis, usia, jenis kelamin, atau faktor risiko lainnya. Setelah data diproses dan dipilah, algoritma KNN dapat digunakan untuk mengklasifikasikan data berdasarkan data pelatihan dan memberikan prediksi tingkat kesehatan mental seseorang. Hal ini dapat membantu dokter atau profesional kesehatan mental dalam

menentukan diagnosis dan perawatan yang tepat untuk pasien (Hajihashemi et al., 2019).

Namun, perlu diingat bahwa hasil prediksi dari algoritma KNN tidak selalu akurat dan dapat dipengaruhi oleh banyak faktor. Oleh karena itu, metode ini perlu dipadukan dengan penilaian medis dan analisis lebih lanjut untuk menghasilkan diagnosis yang akurat dan perawatan yang tepat.

### **2.3.3.2 KNN WOA**

KNN dan WOA (Whale Optimization Algorithm) adalah dua algoritma yang berbeda dalam dunia machine learning dan optimasi. Namun, keduanya dapat digunakan bersama-sama untuk meningkatkan performa prediksi dan optimasi.

KNN dapat digunakan sebagai algoritma klasifikasi atau regresi pada data yang kompleks dan heterogen. Sedangkan WOA adalah algoritma optimasi yang meniru perilaku kelompok paus dalam mencari mangsa di lautan. Algoritma ini bekerja dengan menyebar paus ke seluruh lautan untuk mencari mangsa dan berkumpul kembali untuk melakukan serangan Bersama (Hajihashemi et al., 2019).

KNN dapat digunakan sebagai algoritma klasifikasi atau regresi pada data yang kompleks dan heterogen. KNN cocok digunakan pada data yang tidak terstruktur atau tidak memiliki pola yang jelas. Dalam hal klasifikasi, KNN mengklasifikasikan titik data baru ke kelas mayoritas dari k titik data terdekat. Sedangkan dalam regresi, KNN menghitung nilai rata-rata atau median dari k titik data terdekat untuk memprediksi nilai kontinu.

Sementara itu, algoritma WOA adalah algoritma optimasi yang terinspirasi dari perilaku kelompok paus dalam mencari mangsa di lautan. Algoritma ini bekerja dengan menyebar paus ke seluruh lautan untuk mencari mangsa dan berkumpul kembali untuk melakukan serangan bersama. Pada setiap iterasi, posisi paus diperbarui dengan menggabungkan taktik eksplorasi dan eksploitasi untuk menemukan solusi optimal, WOA telah diterapkan pada berbagai masalah optimasi, termasuk optimasi fungsi, perutean, dan pembelajaran mesin.

Salah satu cara untuk memadukan KNN dan WOA adalah dengan menggunakan WOA untuk memperbaiki atau mengoptimalkan parameter  $k$  pada KNN. Parameter  $k$  pada KNN menentukan jumlah titik data terdekat yang akan diambil untuk menghitung kelas atau nilai regresi. Penulisan WOA dalam optimasi parameter  $k$  dapat meningkatkan akurasi prediksi pada KNN dengan cara mengurangi jumlah iterasi yang dibutuhkan. Semakin sedikit jumlah iterasi, maka semakin mendekati real hasil diagnosis.

Langkah-langkah dalam mengimplementasikan perpaduan KNN dan WOA adalah sebagai berikut:

- (1) Inisialisasi populasi paus WOA.
- (2) Hitung jarak antara data dan tetangga terdekat menggunakan nilai  $k$  yang sudah ditentukan.
- (3) Hitung performa prediksi KNN menggunakan parameter  $k$  yang ditentukan.
- (4) Hitung nilai fungsi objektif yang akan dioptimasi, dalam hal ini adalah nilai performa prediksi KNN.
- (5) Gunakan algoritma WOA untuk mengoptimalkan parameter  $k$  pada KNN.



Ulangi langkah 2-5 untuk seluruh populasi paus WOA.

Pilih posisi paus (best position) dengan performa prediksi KNN terbaik sebagai hasil akhir.

Dengan memadukan KNN dan WOA, diharapkan dapat meningkatkan akurasi prediksi dan efisiensi optimasi pada data yang kompleks dan heterogen. Namun, perlu diingat bahwa perpaduan ini juga memerlukan waktu dan sumber daya yang cukup untuk proses optimasi yang lebih kompleks.



## **BAB III**

### **METODE PENELITIAN**

#### **3.1. Jenis, Sifat, dan Pendekatan Penelitian**

Penulis menggunakan jenis penelitian eksperimental. Hal ini dilakukan karena terdapat serangkaian tindakan dimulai dari penentuan variable uji pada mental health, penyusunan tingkat kecocokan untuk setiap variabel, susun code berdasarkan algoritma WOA (inisialisasi populasi variable yang ditentukan beserta posisi awal agent dan juga posisi akhir agent yang di inginkan, menentukan iterasi maksimal, eksplorasi posisi, penentuan posisi agent selanjutnya, pengujian kecocokan posisi agent, iterasi bila diperlukan), proses olah data training dan proses olah data testing dengan KNN, membandingkan hasil dengan algoritma hibrida KNN – WOA dan algoritma KNN – Edge Enhancement WOA.

Kemudian sifat penelitian ini dilakukan secara mandiri dengan metode deskriptif dari data yang diperoleh kemudian di training dan di testing dengan data dari referensi yang valid, kemudian digunakan algoritma lain sebagai pembanding. Untuk pendekatan, digunakan pendekatan kuantitatif dari data yang diperoleh kemudian di training dan di validasi dengan data yang diujikan kemudian. Kemudian dicari nilai akurasi dengan metode confusion matrix dalam melakukan diagnostik.

### 3.2. Metode Pengumpulan Data

Data di ambil dari dataset penelitian sebelumnya yakni Chung, J. & Teo, J. 2022. Dataset berisi indikator mental health dan history ini telah digunakan untuk mendiagnostik mental health namun dengan non-natural algorithm.

### 3.3. Metode Analisis Data

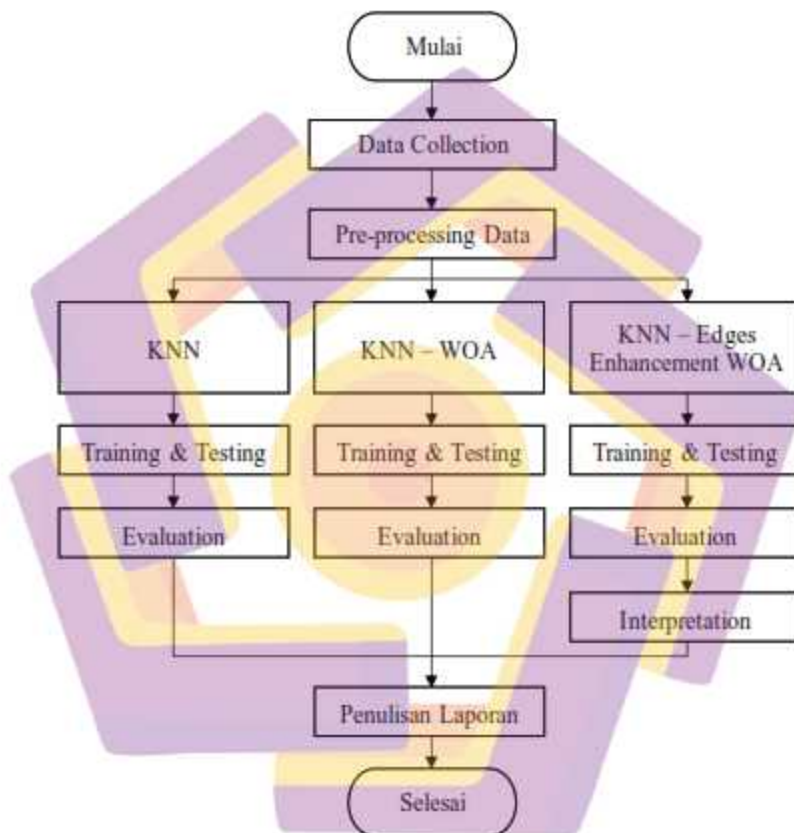
Data naratif akan disajikan kedalam bentuk deskriptif. Sedangkan data kuantitatif akan diolah dengan menggunakan Bahasa Python di Google Colab edisi free. Hasilnya akan disajikan dalam bentuk tabel, grafik, dan uraian secara narasi. Besarnya nilai akurasi akan ditentukan dengan bantuan confusion matrix. Metode ini adalah metode pengukuran performa untuk masalah klasifikasi machine learning. Keluaran digunakan untuk menghitung akurasi.

Penentuan faktor yang dominan mempengaruhi time processing dilakukan dengan analisa visual deskriptif terhadap process ketika running code. Hal ini dilakukan terkait dengan tujuan awal penulis untuk mengetahui faktor apa saja yang secara dominan mempengaruhi time processing. Faktor tersebut akan penulis paparkan pada kolom saran penelitian selanjutnya dalam mengantisipasi faktor tersebut.

Kemudian analisis data juga dilakukan dengan analisa standar deviasi. Standar deviasi adalah ukuran statistik yang mengukur seberapa jauh nilai-nilai data tersebar dari rata-rata. Semakin tinggi standar deviasi, semakin besar variasi atau perbedaan antara nilai-nilai data. Semakin rendah standar deviasi, semakin sedikit variasi atau perbedaan antara nilai-nilai data.

### 3.4. Alur Penelitian

Penelitian ini menggunakan desain penelitian action research yang diimplementasikan ke alur penelitian seperti ditunjukkan pada Gambar 3.1.

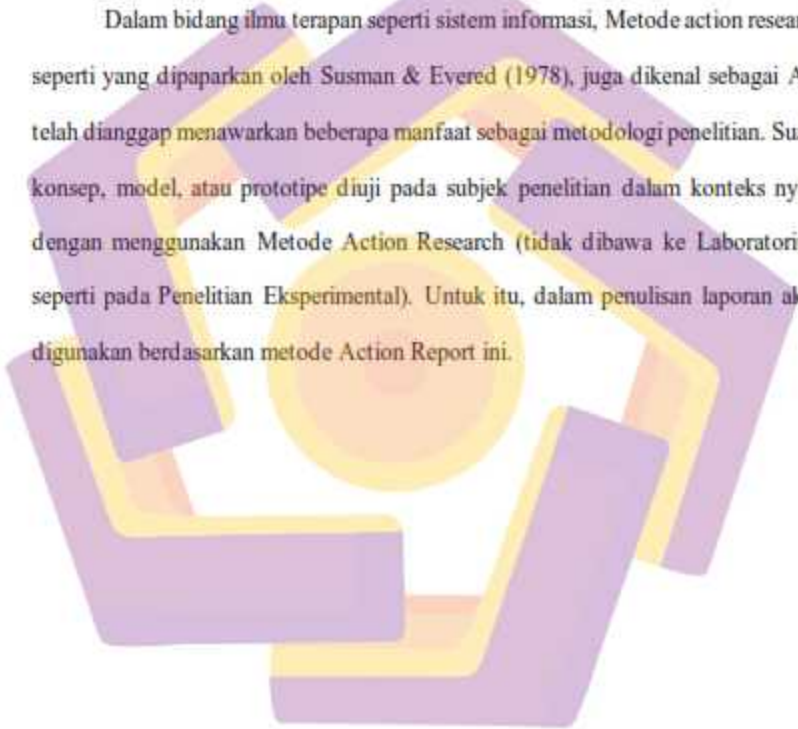


Gambar 3.1 Alur Penelitian

Letak perbedaan KNN, KNN-WOA, dan KNN-Edge Enhancement WOA terletak dari penentuan nilai  $K$  pada algoritma KNN. Nilai  $K$  pada KNN ditentukan berdasarkan Dunia Praktis berfungsi sebagai "Laboratorium" dalam metodologi

Action research untuk menguji teori. Dianggap bahwa mempelajari bagaimana teknologi diterapkan di bidang sistem informasi tidak boleh terbatas pada pemeriksaan teknologi, orang, atau proses saja. Penelitian ini mencakup proses siklus lima fase yang digunakan yaitu Diagnosing, Action Planning, Action Taking, Evaluation, serta ditambahkan dengan Report di akhir.

Dalam bidang ilmu terapan seperti sistem informasi, Metode action research seperti yang dipaparkan oleh Susman & Evered (1978), juga dikenal sebagai AR, telah dianggap menawarkan beberapa manfaat sebagai metodologi penelitian. Suatu konsep, model, atau prototipe diuji pada subjek penelitian dalam konteks nyata dengan menggunakan Metode Action Research (tidak dibawa ke Laboratorium seperti pada Penelitian Eksperimental). Untuk itu, dalam penulisan laporan akan digunakan berdasarkan metode Action Report ini.



## BAB IV

### HASIL PENELITIAN DAN PEMBAHASAN

#### 4.1 Prosedur Baru

##### 4.1.1 Perumusan

Dalam pemaparan Persamaan (1) dan Persamaan (2), iterasi saat ini dilambangkan dengan  $t$  dan iterasi selanjutnya dilambangkan dengan  $t + 1$ . Posisi Best Whale (BW) pada iterasi  $t$  diwakili oleh  $X$ , dan posisi paus tertentu pada iterasi  $t$  diwakili oleh  $X_i$ .

Setelah dilakukan perhitungan pada iterasi  $t + 1$  untuk mendapatkan posisi  $A$  dan  $C$ , langkah selanjutnya adalah melakukan evaluasi nilai fungsi objektif pada masing-masing posisi tersebut. Jika nilai fungsi objektif pada posisi  $A$  lebih baik daripada nilai fungsi objektif pada posisi  $X_i$ , maka posisi  $X_i$  akan diperbarui dengan posisi  $A$ . Begitu pula, jika nilai fungsi objektif pada posisi  $C$  lebih baik daripada nilai fungsi objektif pada posisi BW, maka posisi BW akan diperbarui dengan posisi  $C$ .

Untuk menghitung posisi  $A$  pada iterasi  $t+1$ , digunakan persamaan berikut:

$$A = 2 * r1 * X - X_i \dots\dots\dots(9)$$

dimana  $r1$  adalah bilangan acak antara 0 dan 1.

Setelah melakukan pembaruan posisi, langkah selanjutnya adalah melakukan iterasi berikutnya dengan menggunakan posisi BW dan posisi paus lainnya yang telah diperbarui sebelumnya. Iterasi ini dilakukan hingga mencapai

kondisi berhenti yang telah ditentukan sebelumnya, seperti mencapai batas iterasi maksimum atau mencapai nilai fungsi objektif yang memenuhi kriteria tertentu.

Untuk menghitung posisi C pada iterasi t+1, digunakan persamaan berikut:

$$C = 2 * r2 * BW - X \dots\dots\dots(10)$$

dimana r2 adalah bilangan acak antara 0 dan 1.

Dengan demikian, WOA dapat digunakan untuk menyelesaikan berbagai jenis permasalahan optimasi dengan mengoptimalkan pengurangan jumlah iterasi yang diperlukan dalam melalui proses perburuan kolektif paus bungkuk.

Dalam strategi perburuan acak yang dijelaskan oleh Mirjalili dan Lewis seperti Persamaan (1) hingga Persamaan (8), setiap paus mencari mangsa dengan memperbarui posisinya secara acak. Untuk melakukan ini, mereka menggunakan persamaan berikut untuk menghitung perubahan posisi paus:

$$x(t+1) = x(t) + D \times r \dots\dots\dots(11)$$

dengan:

x(t) adalah posisi paus pada iterasi ke-t.

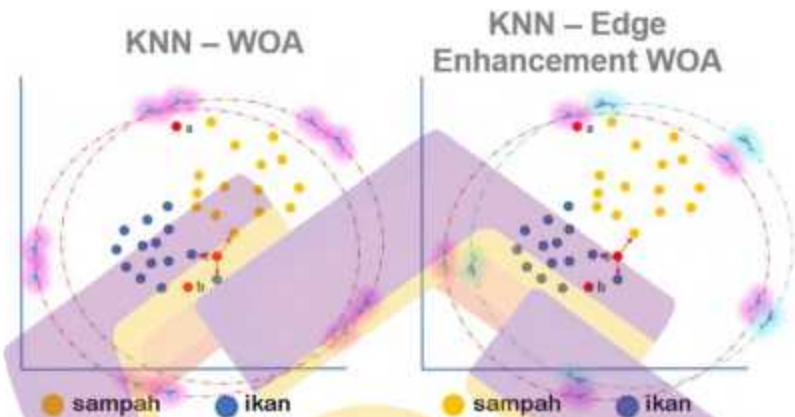
x(t+1) adalah posisi paus pada iterasi ke-(t+1), setelah melakukan perubahan posisi.

D adalah faktor skala yang digunakan untuk mengatur ukuran perubahan posisi.

r adalah vektor perubahan posisi yang diambil secara acak dari distribusi normal dengan nilai  $r \in [-1,1]$ .

Perlu diketahui bahwa KNN-EEWOA tidak memperlakukan perubahan lintasan ketika dua lintasan agent (paus) mengalami intersection seperti pada

Gambar 4.1 ketika kita menganalogikannya dengan ikan paus (agent), ikan (variabel 1), dan sampah (variabel 2).



Gambar 4.1 Penentuan Lintasan Batas Berdasarkan Intersection

Ketika dirumuskan, maka lintasan KNN-EEWOA akan tetap pada lintasanya masing-masing sesuai Persamaan (12) dan Persamaan (13).

Ketika dua garis lintasan lingkaran bersinggungan namun tetap pada lintasanya, maka persamaan matematika yang menggambarkan kondisi ini dapat dinyatakan sebagai berikut:

$$(x - a1)^2 + (y - b1)^2 = r1^2 \dots\dots\dots (12)$$

$$(x - a2)^2 + (y - b2)^2 = r2^2 \dots\dots\dots (13)$$

Dimana  $a1$ ,  $b1$ ,  $r1$  adalah koordinat pusat dan jari-jari lingkaran pertama, sedangkan  $a2$ ,  $b2$ ,  $r2$  adalah koordinat pusat dan jari-jari lingkaran kedua.

Dalam strategi perburuan acak ini, setiap paus memilih posisi baru secara acak pada setiap iterasi. Posisi baru ini ditentukan dengan menambahkan perubahan posisi yang dihasilkan dari perkalian faktor skala  $D$  dengan vektor perubahan posisi



acak  $r$ . Nilai  $r$  diambil dari distribusi normal, sehingga menghasilkan perubahan posisi yang acak dan tidak terduga. Faktor skala  $D$  digunakan untuk mengatur ukuran perubahan posisi, dan dapat disesuaikan untuk mengoptimalkan strategi perburuan paus dalam menemukan mangsa.

Pada perhitungan dengan menggunakan Ms Excel, terdiri 1259 data. Dari keseluruhan data, 1000 data digunakan sebagai data training (data ke-1 sampai data ke-1000) dan 50 data digunakan sebagai data testing (data ke-1001 sampai data ke-1050). Dengan jumlah pilihan 3 (Yes, No, Maybe) dan jumlah agent (paus) sebanyak 2 (default), maka didapatkan 1 lintasan agent sehingga kluster pada iterasi pertama ( $t$ ) ditentukan dengan nilai 1 seperti pada Gambar 4.2. Setelah itu, iterasi kedua ( $t+1$ ) dilakukan dengan kluster dengan nilai 2. Telah dilakukan iterasi dua kali dan nilai  $k$  masih berada pada dua pilihan, yakni  $k=7$  dan  $k=9$ . Jadi masih perlu dilakukan iterasi hingga nilai  $k$  terbaik ditemukan. Pada Gambar 4.2, terlihat bahwa pada kasus dengan intersection, batas-batas kluster (gelembung) otomatis bernilai  $x = 2$  dan  $x = 14$ .



Gambar 4.2 KNN-WOA

Metode KNN-EEWOA tidak melakukan intersection sehingga pada Gambar 4.3 terlihat bahwa ketika intersection tidak dilakukan, dengan kata lain

batas gelembung ditentukan sejak awal maka nilai kluster pada iterasi pertama ( $t$ ) adalah 2. Melanjutkan dari penjelasan sebelumnya, pada metode KNN-EEWOA, ketika tidak dilakukan intersection maka nilai kluster pada iterasi pertama ( $t$ ) akan ditentukan sejak awal. Pada Gambar 4.3, terlihat bahwa pada kasus tanpa intersection, batas-batas kluster (gelembung) ditentukan sejak awal, yaitu pada nilai  $x = 2$  dan  $x = 12$  dan juga nilai  $x = 4$  dan  $x = 14$ . Dengan demikian, ketika dilakukan proses clustering pada iterasi pertama, setiap data point akan dikelompokkan ke dalam kluster yang memiliki batas-batas tersebut. Sehingga, pada kasus ini nilai kluster pada iterasi pertama ( $t$ ) adalah 2, karena terdapat 2 kluster yang dibentuk oleh batas-batas tersebut.



Gambar 4.3 KNN-EEWOA

Kemudian didapatkan nilai kluster pada iterasi kedua sebesar 1, dan alhasil nilai  $k$  terbaik didapatkan tanpa perlu dilakukan iterasi berlanjut. Jika nilai kluster pada iterasi kedua sudah sama dengan nilai kluster pada iterasi sebelumnya, maka dapat disimpulkan bahwa hasil klusterisasi sudah stabil dan tidak perlu dilakukan iterasi berlanjut. Dalam hal ini, nilai  $k$  terbaik dapat diambil dari iterasi tersebut.

Tindakan ini memerlukan pengaturan inisialisasi awal sehingga jumlah kluster pada iterasi pertama tidak dipengaruhi oleh intersection secara otomatis. KNN-EEWOA memerlukan pengaturan diawal penggunaannya.

#### 4.1.2 Langkah Implementasi dalam Komputasi

Penulisan WOA (Whale Optimization Algorithm) dalam optimasi parameter  $k$  pada KNN dapat meningkatkan akurasi prediksi pada KNN. WOA dapat digunakan untuk memperbaiki atau mengoptimalkan nilai parameter  $k$  dengan mencari nilai  $k$  yang memberikan akurasi tertinggi pada data yang diberikan. Hal ini dapat disebut dengan KNN-EEWOA.

Langkah-langkah umum untuk menggabungkan KNN dan EEWOA dalam optimasi parameter  $k$  adalah sebagai berikut:

- 1) Bagi data menjadi dua set: training set dan test set.
- 2) Lakukan preprocessing data untuk menghilangkan nilai-nilai yang hilang atau data yang tidak relevan serta standarisasi nilai numerik agar tidak mempengaruhi hasil klasifikasi atau regresi.
- 3) Hitung jarak antara titik data menggunakan metrik jarak yang sesuai, seperti Euclidean distance untuk data numerik atau metrik jarak Hamming untuk data kategorikal biner.
- 4) Tentukan nilai parameter  $k$  yang akan dioptimalkan. Gunakan nilai  $k$  dari  $k = 1$  hingga  $k = n$  ketika akurasi terus menurun.
- 5) Bangkitkan populasi awal paus secara acak dan hitung fitness dari setiap paus. Dikarenakan KNN-EEWOA tidak berfokus pada jumlah agent (ikan paus), maka digunakan nilai default 2. Jumlah pilihan tergantung seberapa banyak opsi yang digunakan.
- 6) Untuk setiap iterasi, lakukan operasi penyebaran (spreading) paus ke seluruh lautan untuk mencari solusi yang lebih baik. Jangan lakukan

intersection, namun berfokus pada lintasan yang telah ditentukan sebelumnya seperti pada persamaan. Selama operasi penyebaran, perbarui posisi setiap paus menggunakan persamaan WOA yang relevan.

- 7) Jika paus terbaik baru lebih baik dari paus terbaik sebelumnya, gunakan nilai parameter  $k$  dari paus terbaik baru.
- 8) Lakukan klasifikasi atau regresi pada data test dengan menggunakan KNN dengan nilai  $k$  yang dioptimalkan dari hasil WOA.
- 9) Tentukan nilai kluster pada WOA tanpa melakukan intersection. Hal ini dilakukan dengan penerapan Persamaan 12 dan Persamaan 13.
- 10) Dengan penerapan ini, jumlah kluster disesuaikan dengan Batasan bubble yang ada.
- 11) Lakukan terus hingga iterasi menunjukkan satu nilai  $k$  dengan nilai akurasi terbaik.

Dengan menggabungkan KNN dan EEWOA dalam optimasi parameter  $k$ , dapat meningkatkan akurasi prediksi pada KNN dan mengurangi jumlah iterasi yang dibutuhkan. Serta tingkat akurasi menurun, dengan kata lain kesalahan dalam klasifikasi atau regresi pada data yang kompleks dan heterogen akan menurun.

## **4.2 Action Planning**

### **4.2.1 Pengumpulan Data**

Data yang digunakan ialah data open sourcing dan bersifat publik di internet. Data hasil survey ini dilakukan oleh Open Sourcing Mental Illness (OSMI). OSMI melakukan survey pada pekerja di bidang teknologi tentang

kesehatan mental dan sikap terhadap kesehatan mental di tempat kerja. Data dari OSMI seperti pada Tabel 4.1 yang menunjukkan 5 data awal seharusnya dapat memberikan gambaran tentang bagaimana kesehatan mental diperhatikan di industri teknologi dan sikap perusahaan terhadap masalah ini. Data yang digunakan ialah data open sourcing dan bersifat publik di internet. Data hasil survey ini dilakukan oleh Open Sourcing Mental Illness (OSMI), sebuah organisasi nirlaba yang berfokus pada isu kesehatan mental di dunia teknologi. OSMI melakukan survey pada pekerja di bidang teknologi tentang kesehatan mental dan sikap terhadap kesehatan mental di tempat kerja. Data dari OSMI seperti pada Tabel 4.1 menunjukkan informasi-informasi seperti jenis kelamin responden, usia, negara tempat bekerja, status pekerjaan, apakah pernah didiagnosis mengalami masalah kesehatan mental, apakah pernah mencari bantuan untuk kesehatan mental, dan sikap perusahaan terhadap masalah kesehatan mental. Dengan menggunakan data tersebut, dapat dianalisis bagaimana kesehatan mental diperhatikan di industri teknologi dan sikap perusahaan terhadap masalah ini.

Tabel 4.1 Dataset Sebelum di Olah

| Timestamp       | Age | Gender | Country        | state | self_employed | ... |
|-----------------|-----|--------|----------------|-------|---------------|-----|
| 8/27/2014 11:29 | 37  | Female | United States  | IL    | NA            |     |
| 8/27/2014 11:29 | 44  | M      | United States  | IN    | NA            |     |
| 8/27/2014 11:29 | 32  | Male   | Canada         | NA    | NA            |     |
| 8/27/2014 11:29 | 31  | Male   | United Kingdom | NA    | NA            |     |
| 8/27/2014 11:30 | 31  | Male   | United States  | TX    | NA            |     |

Tabel 4.1 (Lanjutan) Dataset Sebelum di Olah

| ... | treatment | work interfere | no employees   | remote work | ... |
|-----|-----------|----------------|----------------|-------------|-----|
|     | Yes       | Often          | 25-Jun         | No          |     |
|     | No        | Rarely         | More than 1000 | No          |     |
|     | No        | Rarely         | 25-Jun         | No          |     |
|     | Yes       | Often          | 26-100         | No          |     |
|     | No        | Never          | 100-500        | Yes         |     |

| ... | tech company | benefits   | care options | wellness program | ... |
|-----|--------------|------------|--------------|------------------|-----|
|     | Yes          | Yes        | Not sure     | No               |     |
|     | No           | Don't know | No           | Don't know       |     |
|     | Yes          | No         | No           | No               |     |
|     | Yes          | No         | Yes          | No               |     |
|     | Yes          | Yes        | No           | Don't know       |     |

| ... | seek help  | anonymity  | leave              | mental health consequence | ... |
|-----|------------|------------|--------------------|---------------------------|-----|
|     | Yes        | Yes        | Somewhat easy      | No                        |     |
|     | Don't know | Don't know | Don't know         | Maybe                     |     |
|     | No         | Don't know | Somewhat difficult | No                        |     |
|     | No         | No         | Somewhat difficult | Yes                       |     |
|     | Don't know | Don't know | Don't know         | No                        |     |

| ... | phys health consequence | coworkers    | supervisor | mental health interview | ... |
|-----|-------------------------|--------------|------------|-------------------------|-----|
|     | No                      | Some of them | Yes        | No                      |     |
|     | No                      | No           | No         | No                      |     |
|     | No                      | Yes          | Yes        | Yes                     |     |
|     | Yes                     | Some of them | No         | Maybe                   |     |
|     | No                      | Some of them | Yes        | Yes                     |     |

Tabel 4.1 (Lanjutan) Dataset Sebelum di Olah

| ... | phys health interview | mental vs physical | obs consequence | comments | ... |
|-----|-----------------------|--------------------|-----------------|----------|-----|
|     | Maybe                 | Yes                | No              | NA       |     |
|     | No                    | Don't know         | No              | NA       |     |
|     | Yes                   | No                 | No              | NA       |     |
|     | Maybe                 | No                 | Yes             | NA       |     |
|     | Yes                   | Don't know         | No              | NA       |     |

Namun, perlu diingat bahwa data dari survei tersebut memiliki batasan dan bias tertentu, tergantung pada sampel yang diambil dan cara pengumpulan data yang dilakukan. Dengan kata lain, dilakukan preprocessing data sebelum pengolahan data tersebut dimulai.

Preprocessing data sangat penting untuk menghasilkan analisis yang akurat dan tepat. Beberapa teknik preprocessing data yang dapat dilakukan pada data survei OSMI ini antara lain:

1) Menangani nilai kosong atau hilang (missing values)

Jika terdapat data yang kosong atau hilang, hal ini dapat mempengaruhi hasil analisis. Oleh karena itu, perlu dilakukan penghapusan atau penggantian nilai kosong tersebut. Salah satu teknik penggantian nilai kosong adalah dengan menggunakan nilai rata-rata atau median.

2) Mengatasi data yang tidak relevan atau tidak dibutuhkan

Pada dataset survei, terkadang terdapat atribut atau kolom yang tidak dibutuhkan atau tidak relevan dengan analisis yang dilakukan. Oleh karena itu, perlu dilakukan penghapusan atribut atau kolom tersebut agar tidak mempengaruhi hasil analisis.

### 3) Normalisasi data numerik

Jika dataset mengandung data numerik yang tidak terstandarisasi, perlu dilakukan normalisasi data tersebut agar tidak mempengaruhi hasil analisis. Normalisasi data numerik dapat dilakukan dengan menggunakan teknik seperti Min-Max Scaling atau Z-Score Scaling.

### 4) Encoding data kategorikal

Jika dataset mengandung data kategorikal, perlu dilakukan encoding agar dapat diolah oleh algoritma KNN. Encoding dapat dilakukan dengan menggunakan teknik seperti One-Hot Encoding atau Label Encoding.

Dengan melakukan preprocessing data yang tepat, hasil analisis dapat menjadi lebih akurat dan dapat memberikan gambaran yang lebih jelas tentang kesehatan mental di industri teknologi dan sikap perusahaan terhadap masalah ini.

#### 4.2.2 Preprocessing Data

Preprocessing data dalam penelitian ini sama dengan preprocessing pada umumnya berupa proses untuk membersihkan, mengubah, dan menyesuaikan data mentah menjadi format yang lebih mudah diolah dan dapat memberikan hasil analisis yang lebih akurat. Beberapa tahapan yang umum dilakukan dalam preprocessing data OSMI antara lain:

- (1) Data transformation, dilakukan tindakan berupa mengubah data mentah menjadi format yang lebih mudah diolah. Tindakan sangat awal berupa mengubah semua jenis huruf menjadi *lowercase*. Hal ini dilakukan untuk memudahkan pengolahan data serta mengurangi *error* pada bahasa



pemrograman yang digunakan. Kemudian dilakukan labeling data dimulai dari 0 hingga 1 untuk dua variable dan 0 hingga 2 untuk tiga variable berlaku seterusnya. Hal ini dilakukan karena pada penelitian ini digunakan algoritma utama yakni KNN yang menggunakan angka sebagai basis utamanya.

- (2) Data cleaning, dilakukan tindakan berupa menghapus atau memperbaiki data yang tidak lengkap, tidak valid, atau duplikat. Pada penelitian ini dilakukan penyesuaian seperti variable *Gender* yang terlalu bervariasi (melebihi dua macam opsional). Dilakukan penyetaraan sehingga hanya tersisa dua opsi yakni *Male* dan *Female* saja. Kemudian dilakukan penyesuaian variable *no\_employees* atau jumlah pekerja yang semula terdiri tidak hanya angka, namun juga angka, bulan, dan kata-kata. Sehingga tersisa opsi berupa angka saja berupa 1-5, 5-25, 26-100, 100-500, 500-1000, dan >1000.
- (3) Data reduction, dilakukan tindakan untuk mengurangi dimensi data dengan cara memilih variabel yang paling relevan dan berdampak besar dalam analisis, serta menghilangkan variabel yang kurang relevan atau redundan. Hal ini dilakukan untuk mempercepat waktu pengolahan data dan menghindari overfitting pada model yang digunakan. Pada tahap ini, dilakukan tindakan berupa pengurangan dimensi data agar lebih mudah diolah dan mengurangi kecacatan model, seperti melakukan seleksi fitur pada *Country*, *state*, serta *comments*.
- (4) Setelah preprocessing data selesai dilakukan, barulah data siap untuk diolah lebih lanjut dan dilakukan analisis yang sesuai. Setelah preprocessing data selesai dilakukan, tahap selanjutnya adalah melakukan pemodelan data dengan

menggunakan algoritma yang sesuai untuk menemukan pola atau informasi yang berguna dari data tersebut. Dalam penelitian ini, salah satu algoritma yang digunakan adalah KNN. Setelah model ditemukan, dilakukan evaluasi model untuk mengevaluasi keakuratan model dalam melakukan prediksi atau klasifikasi data yang belum diketahui. Evaluasi model dapat dilakukan dengan berbagai metode seperti confusion matrix, akurasi, dan standar deviasi. Selain itu, dilakukan juga interpretasi hasil analisis untuk memahami makna atau implikasi dari hasil yang ditemukan.

Data hasil preprocessing seperti pada Tabel 4.2 mengandung nilai yang cocok dan sesuai digunakan dalam pemrosesan dengan menggunakan algoritma utama KNN dan algoritma pendukungnya WOA.

Tabel 4.2 Data Preprocessing

| agee | gender | self_employed | family_history | treatment | ... |
|------|--------|---------------|----------------|-----------|-----|
| 2    | 0      | 0             | 0              | 1         |     |
| 2    | 1      | 0             | 0              | 0         |     |
| 2    | 1      | 0             | 0              | 0         |     |
| 2    | 1      | 0             | 1              | 1         |     |
| 2    | 1      | 0             | 0              | 0         |     |

| ... | work_interfere | no_employees | remote_work | tech_company | ... |
|-----|----------------|--------------|-------------|--------------|-----|
|     | 1              | 1            | 0           | 1            |     |
|     | 2              | 5            | 0           | 0            |     |
|     | 2              | 1            | 0           | 1            |     |
|     | 1              | 2            | 0           | 1            |     |
|     | 0              | 0            | 1           | 1            |     |

Tabel 4.2 (Lanjutan) Data Preprocessing

| ... | benefits | care_options | wellness_program | seek_help | anonymity | ... |
|-----|----------|--------------|------------------|-----------|-----------|-----|
|     | 2        | 1            | 1                | 2         | 2         |     |
|     | 0        | 0            | 0                | 0         | 0         |     |
|     | 1        | 0            | 1                | 1         | 0         |     |
|     | 1        | 2            | 1                | 1         | 1         |     |
|     | 2        | 0            | 0                | 0         | 0         |     |

| ... | leave | mental_health_consequence | phys_health_consequence | ... |
|-----|-------|---------------------------|-------------------------|-----|
|     | 2     | 1                         | 1                       |     |
|     | 0     | 0                         | 1                       |     |
|     | 1     | 1                         | 1                       |     |
|     | 1     | 2                         | 2                       |     |
|     | 0     | 1                         | 1                       |     |

| ... | coworkers | supervisor | phys_health_interview |
|-----|-----------|------------|-----------------------|
|     | 1         | 2          | 0                     |
|     | 0         | 0          | 1                     |
|     | 2         | 2          | 2                     |
|     | 1         | 0          | 0                     |
|     | 1         | 2          | 2                     |

### 4.3 Processing Data

#### 4.3.1 Google Colab

Dilakukan proses **training** dan **testing** dengan perbandingan 80:20. Teknik ini membagi dataset menjadi dua bagian, yaitu dataset pelatihan (training set) sebesar 80% dan dataset pengujian (testing set) sebesar 20%. Pada tahap training, model dilatih menggunakan dataset pelatihan untuk belajar memprediksi nilai target dari data masukan yang diberikan. Setelah model selesai dilatih, pada tahap testing, model diuji menggunakan dataset pengujian untuk mengukur performanya. Dalam proses ini digunakan berbagai tools.

Proses training dan testing adalah bagian penting dalam pengembangan model machine learning. Dalam tahap training, model akan belajar mempelajari pola dan hubungan antara variabel masukan (input) dan variabel target (output) dari data pelatihan. Pada umumnya, teknik optimasi digunakan dalam proses training untuk meminimalkan error atau kesalahan prediksi model terhadap data pelatihan. Sebagai contoh, algoritma Gradient Descent dapat digunakan untuk menemukan nilai bobot (weight) yang optimal dalam model neural network.

Setelah proses training selesai dilakukan, model kemudian diuji dengan menggunakan data pengujian untuk mengukur performanya. Pada tahap ini, model digunakan untuk memprediksi nilai target dari data masukan yang belum pernah dilihat sebelumnya. Tujuan dari tahap testing adalah untuk memvalidasi keakuratan dan kehandalan model dalam memprediksi nilai target dari data yang baru.

Dalam proses training dan testing, terdapat berbagai tools atau perangkat yang dapat digunakan, seperti *Python* dengan *Pandas*, *Numpy*, *WOA pada mealpy.swarm\_based*, *Problem pada mealpy*, *Label Encoder*, *OneHotEncoder pada sklearn.preprocessing*. Tools ini menyediakan berbagai algoritma dan metode untuk pengolahan data, pemodelan, serta evaluasi performa model. Selain itu, tools tersebut juga menyediakan berbagai fungsi dan fitur yang dapat digunakan untuk mengoptimalkan model dan meningkatkan performanya. Digunakan beberapa library seperti:

#### 4.3.1.1 Pandas

Library populer yang digunakan dalam pengolahan data di Python diimplementasikan kedalam *google colab*. Pandas adalah sebuah library open-source yang sering digunakan untuk melakukan manipulasi, analisis, dan visualisasi data dalam bahasa pemrograman Python. Pandas berfungsi sebagai alat untuk memproses data yang dapat digunakan untuk mengolah berbagai macam data seperti data tabular, time-series, dan data dengan format yang bervariasi. Library ini menyediakan struktur data yang efisien digunakan dalam penelitian ini seperti *DataFrame* dan *Series* yang memungkinkan penulis untuk melakukan manipulasi data dengan mudah. Pandas dalam penelitian ini sangat berguna dalam pengolahan data dan berperan penting dalam implementasi KNN, KNN-WOA, dan KNN Edge Enhancement WOA.

#### 4.3.1.2 Numpy

Library Python ini sangat berguna dalam pengolahan data numerik dalam penelitian ini. Library ini menyediakan struktur data seperti array dan matrix yang efisien dan cepat dalam melakukan operasi matematika pada data numerik. Hal ini sangat mendukung kinerja basis KNN itu sendiri. NumPy (Numerical Python) dikatakan sebagai library yang digunakan untuk melakukan komputasi numerik pada bahasa pemrograman Python. NumPy menyediakan struktur data array multidimensi (*ndarray*) yang efisien dan cepat, serta berbagai fungsi matematika untuk melakukan operasi pada array tersebut.

#### 4.3.1.3 WOA pada `mealpy.swarm_based`

Mealpy adalah sebuah library Python yang menyediakan berbagai implementasi algoritma optimisasi swarm-based untuk pengolahan data dan machine learning. Salah satu algoritma optimisasi yang disediakan oleh Mealpy adalah Whale Optimization Algorithm (WOA), yang dapat digunakan untuk memecahkan masalah optimisasi dalam machine learning.

Library WOA (Whale Optimization Algorithm) pada Python dapat ditemukan dalam `mealpy.swarm_based`. WOA adalah algoritma optimasi yang terinspirasi dari perilaku paus biru saat mencari makanan di laut. Algoritma ini termasuk dalam kelompok algoritma swarm-based, di mana beberapa individu (whale) bergerak secara bersamaan untuk mencari solusi terbaik.

Dalam `mealpy.swarm_based`, terdapat beberapa kelas yang dapat digunakan untuk implementasi algoritma WOA, seperti `WOA()`, `EnhancedWOA()`, dan lain-lain. Masing-masing kelas memiliki parameter yang berbeda-beda, tergantung pada kebutuhan pengguna dalam mengoptimasi suatu masalah.

Untuk menggunakan algoritma WOA dari Mealpy, penulis dapat mengimport WOA dari sub-package `swarm_based`. Dengan mengimport WOA, penulis dapat membuat instansiasi dari kelas WOA dan menentukan parameter yang diperlukan untuk menjalankan algoritma optimisasi. Dengan menggunakan Mealpy dan algoritma WOA, penulis dapat melakukan optimisasi pada model machine learning KNN dan menghasilkan hasil yang lebih baik dengan cepat. Mealpy menyediakan berbagai fungsi dan konfigurasi untuk menyesuaikan algoritma WOA dan mengoptimalkan model sesuai dengan kebutuhan penulis.

#### 4.3.1.4 Problem pada mealpy

Pada library Mealpy, terdapat module problem yang digunakan untuk menentukan fungsi-fungsi matematika yang akan dioptimalkan. Module ini menyediakan berbagai macam fungsi bawaan yang umum digunakan dalam optimasi, seperti fungsi Rosenbrock, fungsi Rastrigin, fungsi Griewank, dan sebagainya.

Selain itu, module ini juga memungkinkan pengguna untuk menentukan fungsi kustom sendiri dengan mengikuti template yang disediakan. Fungsi kustom ini dapat disesuaikan dengan kebutuhan pengguna untuk mengoptimalkan masalah tertentu.

Dalam penggunaannya, module problem ini biasanya diintegrasikan dengan algoritma optimasi yang terdapat pada library Mealpy, seperti PSO, GWO, dan WOA. Dengan demikian, pengguna dapat dengan mudah melakukan optimasi pada fungsi-fungsi matematika yang diinginkan tanpa harus menuliskan kode dari awal.

Dengan mengimport Problem dari mealpy, penulis dapat membuat instansiasi dari kelas tersebut dan menentukan masalah optimisasi yang ingin dipecahkan, seperti fungsi tujuan, batasan variabel, dan sebagainya. Selain itu, kelas Problem juga dapat digunakan untuk memanggil fungsi evaluasi yang akan digunakan oleh algoritma WOA dalam setiap iterasi, sehingga model machine learning dapat dioptimalkan dengan lebih efektif.

#### 4.3.1.5 Label Encoder, OneHotEncoder pada sklearn.preprocessing

Sklearn populer digunakan dalam machine learning. Library ini menyediakan berbagai algoritma machine learning, fungsi utilitas, dan tools untuk pengolahan data, preprocessing, dan analisis data.

Library Label Encoder dan OneHotEncoder pada sklearn.preprocessing adalah library untuk melakukan encoding data kategorikal pada dataset menjadi data numerik agar dapat diproses oleh algoritma machine learning.

Label Encoder digunakan untuk mengubah label kategori menjadi bilangan bulat dengan cara memberikan bilangan unik untuk setiap kategori. Contohnya jika dataset memiliki kolom "warna" dengan nilai "merah", "biru", dan "hijau", maka Label Encoder akan mengubah nilai tersebut menjadi 0, 1, dan 2.

OneHotEncoder digunakan untuk mengubah data kategorikal menjadi vektor biner (0 dan 1) yang merepresentasikan setiap nilai kategori. Contohnya jika dataset memiliki kolom "jenis kelamin" dengan nilai "laki-laki" dan "perempuan", maka OneHotEncoder akan mengubah nilai tersebut menjadi dua kolom yaitu "jenis kelamin\_laki-laki" dan "jenis kelamin\_perempuan" dengan nilai 0 dan 1.

Dengan menggunakan Label Encoder dan OneHotEncoder, data kategorikal pada dataset dapat diubah menjadi data numerik yang dapat diproses oleh algoritma machine learning.

Dalam pengolahan data dalam penelitian ini, sklearn menyediakan fungsi preprocessing LabelEncoder dan OneHotEncoder. Keduanya digunakan untuk mengubah data kategorikal menjadi bentuk numerik yang dapat digunakan oleh model KNN, KNN - WOA, dan KNN - Edge Enhancement WOA.



LabelEncoder digunakan untuk mengubah label atau nilai kategori menjadi angka, di mana setiap kategori seperti salah satunya ialah *Male* menjadi 1 dan *Female* menjadi 2 pada variable *Gender*. Sedangkan OneHotEncoder digunakan untuk mengubah nilai kategori menjadi bentuk vektor biner, di mana setiap kategori akan memiliki vektor biner yang merepresentasikan kategori tersebut.

#### 4.3.2 Microsoft Excel

Setelah mempersiapkan data mentah dalam format yang dapat diimpor ke Microsoft Excel berupa file CSV, pastikan file tersimpan dalam format Excel agar perubahan dapat disimpan.

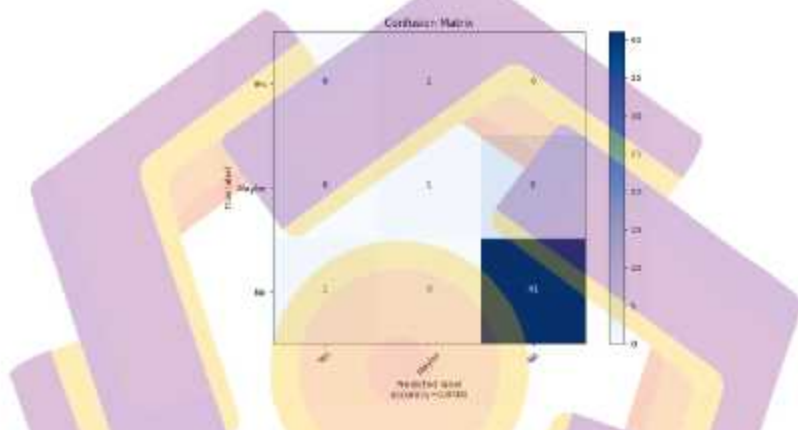
Dari 1258 data keseluruhan, data dibagi menjadi 1000 data training dan 50 data testing. Data training yang ditentukan pada Ms Excel merupakan data dari 1-1000. Data testing yang ditentukan pada Ms Excel merupakan data dari 1001-1050. Atribut untuk setiap sampel data telah ditentukan, atribut *w\_mental\_health\_interview* ditetapkan sebagai atribut penentu.

Digunakan 3 algoritma diantaranya KNN, KNN-WOA, dan KNN-EEWOA. Pada KNN digunakan  $K = 4$  karena opsi atribut *w\_mental\_health\_interview* memiliki 3 opsi: *No*, *Maybe*, dan *Yes*. Pada KNN-WOA, nilai  $K$  dilakukan dengan range 1-50. Hal serupa juga dilakukan pada KNN-EEWOA. Ringkasnya seperti berikut:

- KNN  $\rightarrow k = 4 \rightarrow$  data ke 1001-1050
- KNN-WOA  $\rightarrow k = 1-50 \rightarrow$  data ke 1001-1050
- KNN-EEWOA  $\rightarrow k = 1-50 \rightarrow$  data ke 1001-1050

#### 4.3.2.1 KNN pada $k = 4$

Ketentuan label encoder diserupakan dengan LabelEncoder pada google colab. Kemudian dilakukan pemrosesan data untuk mencari tingkat akurasi dengan data testing yang telah ditentukan sebelumnya. Didapatkan Gambar 4.4 menjelaskan confusion matrix dari hasil pengolahan data pada KNN dengan  $k = 4$ .

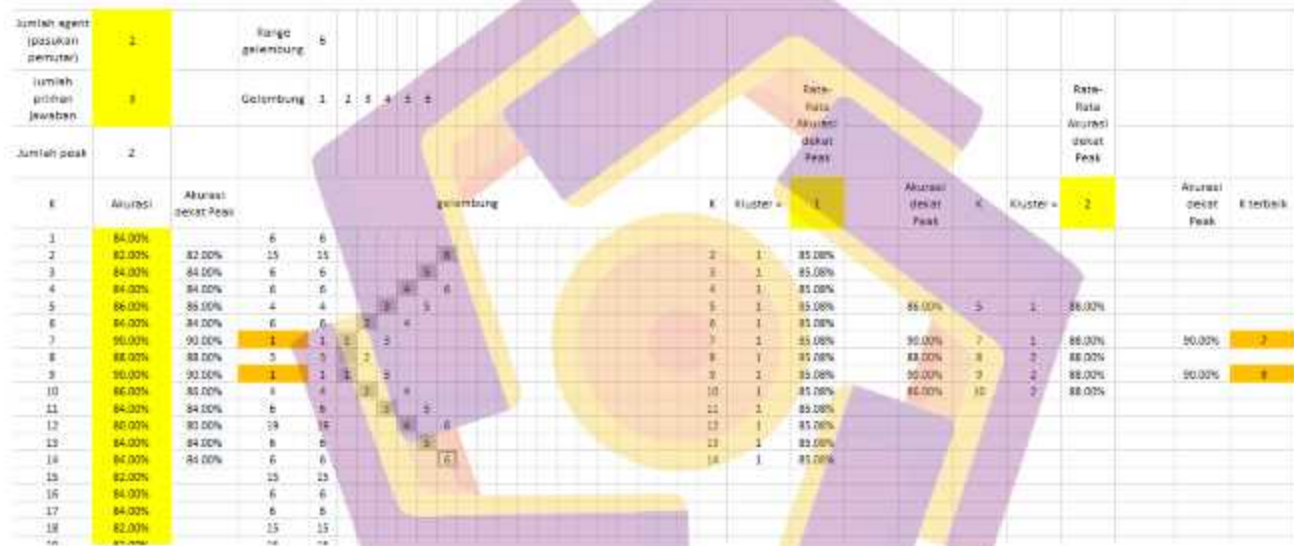


Gambar 4.4 Confusion Matrix data 1001-1050 pada KNN dengan  $k = 4$

Berdasarkan gambar 4.4 perihal confusion matrix KNN  $k = 4$ , didapatkan nilai akurasi, presisi, dan F1-Score berurutan sebesar 84%, 50%, dan 20%.

#### 4.3.2.2 KNN-WOA pada $k = 1-50$

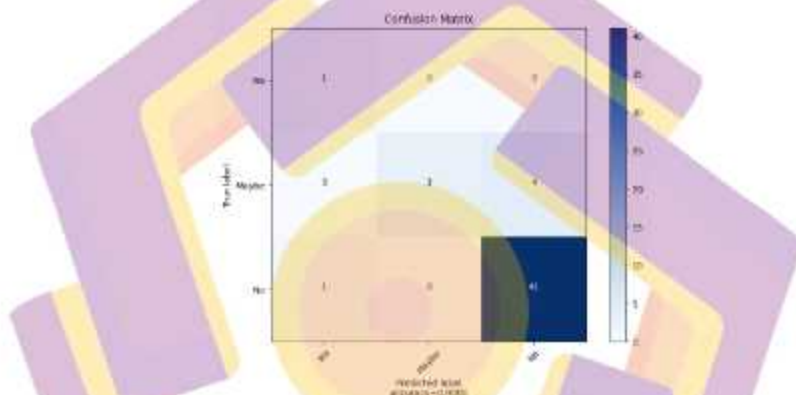
Dengan data serupa dengan data KNN yakni data 1001-1050, digunakan nilai  $k = 1$  hingga  $k = 50$ . Kemudian untuk agent, digunakan agent sebanyak 2 buah atau analoginya 2 ekor ikan paus. Hal ini didasarkan atas default metode untuk atribut dengan opsi sebanyak 3 buah (No, Maybe, Yes). Dengan mengalikan jumlah agent sebanyak 2 ekor dan opsi sebanyak 2 buah, maka gelembung pada metode WOA sebanyak 6 seperti pada Gambar 4.5.



Gambar 4.5 Pemrosesan Data dengan KNN-WOA

Data  $k = 19$  hingga  $k = 50$  tidak menunjukkan tanda-tanda diperlukannya pemberian gelembung. Terlihat terdapat dua peak yakni pada  $k = 7$  dan  $k = 9$ . Metode KNN-WOA secara default akan melakukan pengelompokan data intersection menjadi 1 kluster.

Dikarenakan terdapat 1 kluster saja, maka pada iterasi ke-2 diisikan nilai *rata-rata akurasi dekat peak* sebesar 1 kluster dan didapatkan 2 kelompok kluster berikutnya. Hal serupa tetap terjadi yakni 2 kelompok kluster didapatkan sehingga diperlukan iterasi ke-3 dan seterusnya hingga iterasi ke-7 dimana didapatkan hasil akhir saran penggunaan  $k = 9$  untuk metode KNN-WOA dengan tingkat akurasi mencapai 90% seperti pada Gambar 4.6.

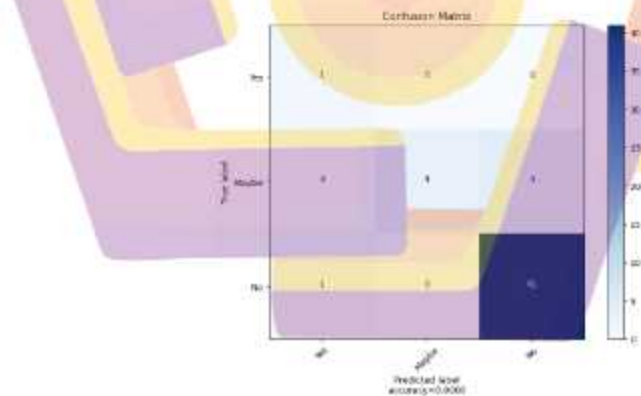


Gambar 4.6 Confusion Matrix data 1001-1050 pada KNN-WOA dengan  $k = 9$

Berdasarkan gambar 4.6 perihal confusion matrix KNN-WOA  $k = 9$ , didapatkan nilai akurasi, presisi, dan F1-Score berurutan sebesar 84%, 80%, dan 61.54%. Dalam hal ini, nilai akurasi menunjukkan seberapa akurat model dalam memprediksi kelas yang benar, sedangkan nilai presisi menunjukkan seberapa baik model dalam mengklasifikasikan data positif. Nilai F1-Score mengkombinasikan nilai akurasi dan presisi untuk memberikan pengukuran yang lebih holistik tentang kinerja model klasifikasi.

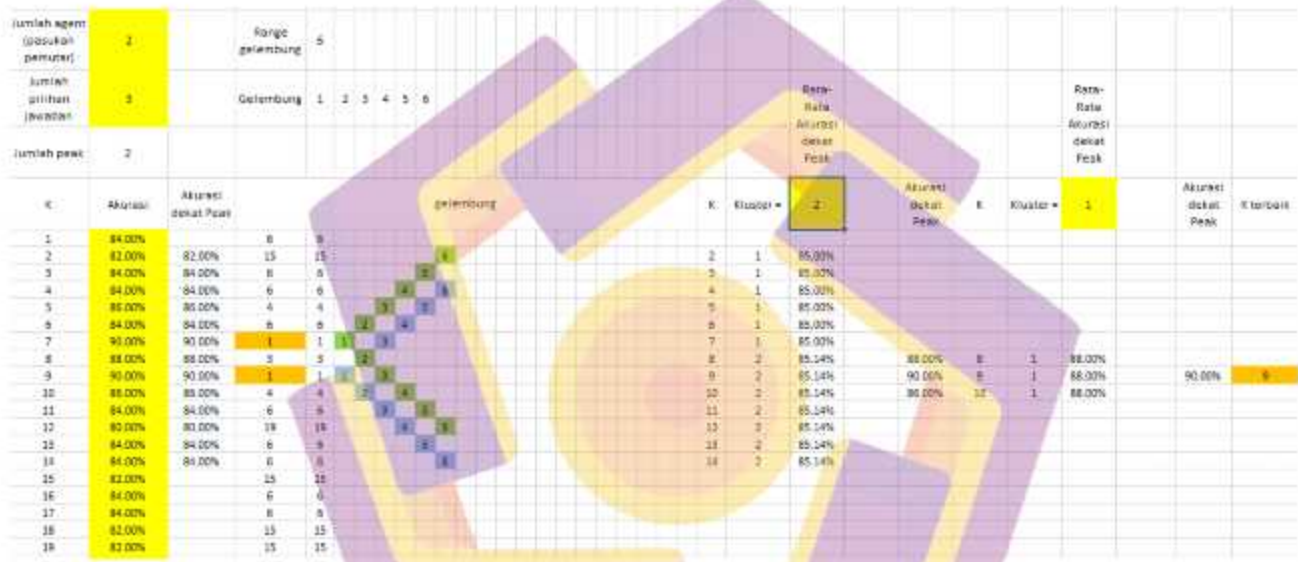
#### 4.3.2.3 KNN-EEWOA pada $k = 1-50$

Dengan data serupa dengan data KNN dan KNN-WOA yakni data 1001-1050, digunakan nilai  $k = 1$  hingga  $k = 50$ . Kemudian untuk agent, digunakan agent sebanyak 2 buah atau analoginya 2 ekor ikan paus sama dengan KNN-WOA sebelumnya. Hal ini didasarkan atas default metode untuk atribut dengan opsi sebanyak 3 buah (No, Maybe, Yes). Dengan mengalikan jumlah agent sebanyak 2 ekor dan opsi sebanyak 2 buah, maka gelembung pada metode WOA sebanyak 6 sama seperti KNN-WOA sebelumnya seperti pada Gambar 4.6. Perbedaan terletak di penentuan batas ujung gelembung dimana intersection tidak mempengaruhi pengambilan keputusan jumlah kluster. Hasilnya meskipun terdapat intersection maka jumlah kluster tetap mengikuti jumlah peak yang muncul seperti pada Gambar 4.5 dengan confusion matrix pada Gambar 4.7.



Gambar 4.7 Confusion Matrix data 1001-1050 pada KNN-EEWOA dengan  $k = 9$

Berdasarkan gambar 4.7 perihal confusion matrix KNN-WOA  $k = 9$ , didapatkan nilai akurasi, presisi, dan F1-Score berurutan sebesar 84%, 80%, dan 61.54%.



Gambar 4.8 Pemrosesan Data dengan KNN-EEWOA

Data k = 20 hingga k = 50 tidak menunjukkan tanda-tanda diperlukannya pemberian gelembung. Terlihat terdapat dua peak yakni pada k = 7 dan k = 9. Metode KNN-EEWOA tidak mengikuti default dimana pengelompokan data intersection menjadi 1 kluster seperti pada Gambar 4.8.

Dikarenakan terdapat 2 kluster, maka pada iterasi ke-2 diisikan nilai *rata-rata akurasi dekat peak* sebesar 2 kluster dan didapatkan 1 kelompok kluster berikutnya. Secara langsung tidak diperlukan lagi iterasi ke-2 karena hasil akhir telah didapatkan yakni saran penggunaan  $k = 9$  untuk metode KNN-EEWOA dengan tingkat akurasi mencapai 90%.

#### 4.3.3 Variabel yang digunakan

Terdapat 22 variabel independent serta 1 variabel dependent. Keseluruhan 22 variabel independent tersebut terdiri dari:

##### *age*

Umur responden, Variabel "Umur responden" adalah variabel numerik yang mengukur usia dari responden pada saat survey dilakukan. Variabel ini dapat digunakan untuk mengevaluasi hubungan antara usia dengan masalah kesehatan mental yang dialami oleh responden, serta untuk menentukan apakah ada perbedaan persepsi atau sikap terhadap kesehatan mental di tempat kerja antara kelompok usia yang berbeda.

##### *gender*

Jenis kelamin responden, Variabel jenis kelamin pada dataset OSMI mengindikasikan jenis kelamin dari responden, terdiri dari dua kategori yaitu "Male", "Female", dan beberapa pilihan lainnya. Variabel jenis kelamin ini merupakan variabel kategorikal atau nominal, karena tidak dapat diurutkan dan tidak memiliki jarak antar nilai. Variabel jenis kelamin ini dapat digunakan sebagai

faktor dalam analisis kesehatan mental responden pada industri teknologi. Dalam analisis, setelah dikelompokkan menjadi "Male" dan "Female", variabel ini dapat diubah menjadi angka menggunakan teknik label encoding atau one-hot encoding.

### *self\_employed*

Apakah responden pekerja mandiri atau tidak (pekerja kantoran), Variabel "self\_employed" dalam dataset tersebut mengindikasikan apakah responden bekerja sebagai pekerja lepas atau bukan. Nilai yang mungkin adalah "Yes" (ya) dan "No" (tidak). Variabel ini penting untuk memahami kondisi pekerja lepas di industri teknologi dan bagaimana kondisi kesehatan mental mereka.

### *family\_history*

Apakah keluarga responden ada riwayat terganggu mental health. Variabel "family\_history" adalah variabel yang mengindikasikan apakah responden memiliki riwayat keluarga yang menderita gangguan kesehatan mental atau tidak. Variabel ini bersifat kategorikal dengan dua nilai yang mungkin, yaitu "Yes" dan "No". Dalam penelitian kesehatan mental, riwayat keluarga merupakan salah satu faktor risiko yang dapat mempengaruhi kemungkinan seseorang untuk mengalami gangguan kesehatan mental. Dengan mengetahui apakah responden memiliki riwayat keluarga yang menderita gangguan kesehatan mental atau tidak, dapat membantu dalam mengidentifikasi faktor risiko dan mengembangkan strategi pencegahan atau intervensi yang tepat.



***treatment***

Apakah responden sedang dalam proses pengobatan terkait masalah Kesehatan. Variabel *treatment* adalah variabel target pada dataset yang menunjukkan apakah responden pernah mendapatkan pengobatan untuk masalah kesehatan mental di tempat kerja atau tidak. Variabel ini bersifat kategorikal biner dengan nilai "Yes" dan "No". Responden yang memilih "Yes" menunjukkan bahwa mereka pernah mendapatkan pengobatan, sedangkan responden yang memilih "No" menunjukkan bahwa mereka belum pernah mendapatkan pengobatan. Variabel ini sangat penting dalam penelitian ini karena merupakan tujuan utama dari analisis yang dilakukan, yaitu untuk memprediksi kemungkinan responden mendapatkan pengobatan berdasarkan variabel lain yang terdapat pada dataset.

***work\_interfere***

Apakah responden mengalami masalah dalam pekerjaan mereka. Variabel "*work\_interfere*" mengacu pada seberapa sering masalah kesehatan mental responden mempengaruhi pekerjaan mereka.

***no\_employees***

Berapa jumlah pekerja dalam perusahaan tempat responden bekerja. Variabel "*no\_employees*" adalah variabel kategorikal ordinal yang mengindikasikan jumlah karyawan dalam perusahaan tempat responden bekerja. Variabel ini dapat digunakan untuk mengevaluasi apakah ukuran perusahaan mempengaruhi kesehatan mental karyawan dan respons perusahaan terhadap isu-isu kesehatan

mental. Variabel ini juga dapat digunakan sebagai faktor untuk memprediksi kemungkinan seorang karyawan memerlukan perawatan kesehatan mental.

#### ***remote\_work***

Apakah responden bekerja secara remote. Variabel "remote\_work" dalam dataset menjelaskan apakah responden bekerja secara remote atau tidak. Variabel ini memiliki dua nilai yaitu "Yes" dan "No". Nilai "Yes" menunjukkan bahwa responden bekerja secara remote sementara nilai "No" menunjukkan bahwa responden tidak bekerja secara remote. Informasi ini dapat memberikan gambaran tentang seberapa besar perusahaan-perusahaan dalam industri teknologi yang memperbolehkan atau memfasilitasi kerja dari jarak jauh bagi karyawannya. Selain itu, variabel ini juga dapat mempengaruhi tingkat stres dan kesejahteraan mental responden. Responden yang bekerja secara remote mungkin menghadapi tantangan yang berbeda dalam menjaga kesehatan mental mereka dibandingkan dengan responden yang bekerja di kantor.

#### ***tech\_company***

Apakah responden bekerja di perusahaan berbasis high teknologi. Variabel "tech\_company" adalah variabel yang mengindikasikan apakah perusahaan tempat responden bekerja bergerak dalam bidang teknologi atau tidak. Variabel ini memiliki dua nilai yaitu "Yes" dan "No". Nilai "Yes" menunjukkan bahwa perusahaan tempat responden bekerja bergerak dalam bidang teknologi, sedangkan nilai "No" menunjukkan sebaliknya. Variabel ini dapat digunakan untuk melihat

seberapa besar proporsi responden yang bekerja di perusahaan teknologi dan juga dapat digunakan untuk melihat perbedaan sikap atau persepsi terhadap kesehatan mental antara responden yang bekerja di perusahaan teknologi dan yang tidak.

### ***benefits***

Apakah perusahaan menyediakan fasilitas penanganan terkait masalah mental health. Variabel *benefits* adalah variabel yang mengukur apakah responden mendapatkan manfaat atau tunjangan dalam kaitannya dengan masalah kesehatan mental di tempat kerja. Responden diminta untuk menilai seberapa besar manfaat atau tunjangan yang mereka terima dalam hal kesehatan mental di tempat kerja, dengan pilihan jawaban berupa "Yes", "No", atau "I don't know". Variabel ini digunakan sebagai salah satu faktor dalam mengukur sikap perusahaan terhadap kesehatan mental di tempat kerja. Semakin banyak perusahaan yang memberikan manfaat atau tunjangan dalam hal kesehatan mental, dapat dianggap bahwa perusahaan tersebut lebih memperhatikan kesejahteraan mental karyawannya. Namun, perlu diingat bahwa tingkat manfaat atau tunjangan yang diberikan oleh perusahaan mungkin bervariasi dan tidak selalu mencerminkan kepedulian yang sebenarnya terhadap kesehatan mental karyawan.

### ***care\_options***

Apakah responden tahu pilihan untuk perawatan kesehatan mental yang disediakan perusahaan? Variabel *care\_options* mengukur apakah perusahaan memberikan opsi untuk merawat kesehatan mental kepada karyawannya.

Responden diminta untuk memilih opsi "ya", "tidak", atau "tidak yakin". Variabel ini dapat memberikan gambaran tentang seberapa serius perusahaan dalam memperhatikan kesehatan mental karyawan dan seberapa terbuka perusahaan terhadap isu kesehatan mental. Hal ini juga dapat memengaruhi karyawan dalam memilih perusahaan yang memberikan perhatian serius terhadap kesehatan mental mereka.

#### ***wellness\_program***

Pemahkah atasan responden membahas kesehatan mental sebagai bagian dari program kesejahteraan responden? Variabel *wellness\_program* mengacu pada pertanyaan apakah perusahaan memberikan program kesejahteraan untuk karyawan dalam rangka meningkatkan kesehatan fisik dan mental. Responden diminta untuk memilih opsi "Yes", "No", atau "Don't know" sebagai jawaban. Variabel ini dapat memberikan informasi tentang seberapa sering perusahaan memberikan program kesejahteraan bagi karyawan. Hal ini dapat menjadi faktor yang penting dalam mempengaruhi kondisi kesehatan mental dan fisik karyawan. Dalam analisis data, variabel *wellness\_program* dapat dijadikan sebagai faktor yang mempengaruhi keputusan responden untuk menerima atau menolak pengobatan. Hal ini dapat membantu dalam memahami apakah program kesejahteraan yang disediakan perusahaan sudah efektif dalam meningkatkan kesehatan karyawan ataukah masih perlu ditingkatkan.

### *seek\_help*

Apakah responden menyediakan informasi perihal mental health dan bagaimana penanganannya. Variabel *seek\_help* pada dataset tersebut merupakan sebuah kolom yang menunjukkan apakah responden merasa terbuka untuk mencari bantuan terkait masalah kesehatan mental. Kolom ini memiliki dua nilai, yaitu "Yes" dan "No". Responden yang menjawab "Yes" menunjukkan bahwa mereka merasa terbuka untuk mencari bantuan, sedangkan responden yang menjawab "No" menunjukkan bahwa mereka tidak merasa terbuka untuk mencari bantuan terkait masalah kesehatan mental. Variabel ini bisa menjadi salah satu indikator tentang tingkat stigma atau kesadaran terhadap pentingnya kesehatan mental di kalangan responden. Responden yang merasa terbuka untuk mencari bantuan dapat menunjukkan bahwa mereka memiliki kesadaran yang lebih tinggi tentang kesehatan mental dan merasa nyaman untuk mencari bantuan jika diperlukan. Sementara itu, responden yang tidak merasa terbuka untuk mencari bantuan mungkin mengalami stigma atau merasa malu untuk mengakui masalah kesehatan mental yang mereka hadapi.

### *anonymity*

Apakah anonimitas responden terlindungi. Variabel *anonymity* adalah variabel yang mengukur seberapa besar responden merasa aman dan nyaman untuk melaporkan masalah kesehatan mental mereka dengan anonimitas yang diberikan oleh perusahaan tempat mereka bekerja. Responden diminta untuk memilih salah satu dari tiga opsi yang tersedia: "Ya", "Tidak", atau "Saya tidak tahu". Jika

responden merasa bahwa anonimitas yang diberikan oleh perusahaan tempat mereka bekerja membuat mereka merasa lebih aman untuk melaporkan masalah kesehatan mental mereka, maka mereka akan memilih opsi "Ya". Sebaliknya, jika mereka merasa bahwa anonimitas tidak penting atau bahkan membuat mereka merasa tidak aman, mereka akan memilih opsi "Tidak". Sedangkan, jika mereka tidak tahu apa itu anonimitas atau bagaimana hal itu mempengaruhi melaporkan masalah kesehatan mental, mereka akan memilih opsi "Saya tidak tahu". Variabel ini dapat memberikan informasi penting bagi perusahaan untuk memperbaiki kebijakan mereka terkait anonimitas bagi karyawan yang melaporkan masalah kesehatan mental.

#### *leave*

Kemudahan responden melakukan penanganan mental health. Variabel "leave" pada dataset mengindikasikan seberapa mudah responden dapat mengambil cuti atau waktu bebas untuk alasan kesehatan mental. Variabel ini memiliki beberapa kategori, yaitu: (1) Very difficult: Sangat sulit untuk mengambil cuti atau waktu bebas untuk alasan kesehatan mental (2) Somewhat difficult: Agak sulit untuk mengambil cuti atau waktu bebas untuk alasan kesehatan mental (3) Don't know: Tidak tahu seberapa mudah untuk mengambil cuti atau waktu bebas untuk alasan kesehatan mental (4) Somewhat easy: Agak mudah untuk mengambil cuti atau waktu bebas untuk alasan kesehatan mental (5) Very easy: Sangat mudah untuk mengambil cuti atau waktu bebas untuk alasan kesehatan mental Variabel ini dapat memberikan informasi tentang seberapa mendukungnya perusahaan dalam hal

kesehatan mental karyawan dan seberapa sulit atau mudah bagi karyawan untuk mengambil waktu luang jika dibutuhkan untuk kesehatan mental mereka.

#### ***mental\_health\_consequence***

Apakah responden memiliki resiko jika melaporkan isu mental health kepada tempat kerja. Variabel *mental\_health\_consequence* adalah variabel yang mengindikasikan apakah seseorang mengalami dampak atau konsekuensi dari masalah kesehatan mental yang dialaminya. Dalam penelitian ini, responden diminta untuk menjawab apakah mereka mengalami konsekuensi negatif dalam pekerjaan mereka akibat dari masalah kesehatan mental yang mereka alami. Jawaban dari responden dapat berupa "yes", "no", atau "maybe". Variabel ini menjadi penting dalam penelitian ini karena dapat memberikan informasi tentang seberapa sering seseorang mengalami konsekuensi negatif di tempat kerja akibat dari masalah kesehatan mental yang mereka alami. Hasil dari analisis variabel ini dapat memberikan pandangan tentang pentingnya peran perusahaan dalam memberikan dukungan dan fasilitas untuk mengatasi masalah kesehatan mental pada karyawan mereka.

#### ***phys\_health\_consequence***

Apakah responden memiliki resiko jika melaporkan isu Kesehatan fisik kepada tempat kerja. Variabel *phys\_health\_consequence* adalah variabel nominal yang menunjukkan apakah responden berpikir bahwa kondisi kesehatan fisik mereka akan mempengaruhi pekerjaan mereka dan apakah ada konsekuensi negatif

akibat mengungkapkan kondisi kesehatan fisik mereka di tempat kerja. Nilai yang mungkin dari variabel ini adalah "Yes", "No", atau "Maybe". Variabel ini digunakan dalam penelitian ini untuk mengevaluasi apakah ada hubungan antara masalah kesehatan mental dan fisik yang dialami oleh responden, dan apakah ada konsekuensi negatif yang mungkin terjadi jika responden mengungkapkan masalah kesehatan mental dan fisik mereka di tempat kerja. Variabel ini dapat memberikan wawasan tentang bagaimana pengungkapan masalah kesehatan mental dan fisik dapat mempengaruhi karir seseorang dan bagaimana perusahaan dapat merespons masalah ini.

#### ***coworkers***

Apakah responden memiliki resiko jika melaporkan isu mental health kepada rekan kerja. Variabel coworkers ini penting untuk dianalisis karena dapat mempengaruhi kesejahteraan mental karyawan di tempat kerja. Jika seorang karyawan merasa nyaman berbicara tentang masalah kesehatan mental dengan rekan kerja mereka, maka mereka mungkin lebih terbuka dan mudah mendapatkan dukungan atau bantuan dari rekan kerja dalam mengatasi masalah tersebut. Sebaliknya, jika seorang karyawan tidak merasa nyaman berbicara tentang masalah kesehatan mental dengan rekan kerja mereka, maka mereka mungkin cenderung menyembunyikan masalah tersebut dan mengalami stres atau tekanan yang lebih tinggi di tempat kerja.



### *supervisor*

Apakah responden menginginkan berdiskusi perihal mental health kepada atasannya. Variabel supervisor merujuk pada pertanyaan dalam survei terkait dengan pendapat responden tentang ketersediaan atau dukungan dari supervisor mereka dalam mengatasi masalah kesehatan mental di tempat kerja. Pada dasarnya, variabel ini mengindikasikan sejauh mana supervisor di tempat kerja memperhatikan dan mendukung kesehatan mental karyawan mereka. Hasil dari variabel supervisor dapat membantu dalam mengevaluasi keberhasilan program-program dukungan kesehatan mental di tempat kerja yang melibatkan supervisor sebagai bagian dari program tersebut.

### *phys\_health\_interview*

Apakah responden mengumbar perihal mental health kepada pekerja junior. Variabel *phys\_health\_interview* adalah variabel nominal yang mengindikasikan apakah responden merasa nyaman untuk membicarakan topik kesehatan mental dengan calon atasan mereka selama wawancara. Variabel ini memiliki tiga nilai kategori yaitu "Yes", "No", dan "Maybe". Variabel ini penting untuk diketahui karena dapat mempengaruhi keputusan seseorang untuk mengakses layanan kesehatan mental di tempat kerja dan untuk membicarakan masalah kesehatan mental mereka secara terbuka dengan atasan mereka. Jika responden merasa nyaman untuk membicarakan topik kesehatan mental dengan atasan mereka, maka mereka mungkin lebih cenderung untuk mencari dan memanfaatkan layanan kesehatan mental di tempat kerja dan mendapatkan dukungan dari atasan mereka.

Namun, jika mereka tidak merasa nyaman membicarakan topik ini dengan atasan mereka, mereka mungkin enggan untuk mencari bantuan dan mendapatkan dukungan yang dibutuhkan.

### ***obs\_consequence***

Apakah responden memiliki rekan yang mengalami mental health dan mendapatkan dampak buruk akibatnya. Variabel "obs\_consequence" adalah variabel kategorikal yang mengindikasikan apakah responden pernah melihat atau mengalami konsekuensi negatif di tempat kerja mereka karena masalah kesehatan mental mereka. Variabel ini memiliki dua nilai, yaitu "Yes" dan "No". Jika nilai variabel adalah "Yes", maka responden telah mengalami konsekuensi negatif di tempat kerja karena masalah kesehatan mental mereka. Sedangkan jika nilai variabel adalah "No", maka responden tidak pernah mengalami konsekuensi negatif di tempat kerja karena masalah kesehatan mental mereka. Variabel ini akan digunakan dalam analisis untuk mengetahui seberapa sering responden mengalami konsekuensi negatif di tempat kerja karena masalah kesehatan mental mereka.

### ***mental\_vs\_physical***

Menurut responden apakah perihal mental health sama berbahayanya dengan perihal Kesehatan fisik. Variabel "mental\_vs\_physical" pada dataset mengindikasikan apakah responden merasa bahwa kesehatan mental sama pentingnya dengan kesehatan fisik. Variabel ini merupakan tipe data kategorikal dengan dua kemungkinan nilai yaitu "Yes" atau "No". Dalam konteks penelitian

kesehatan mental, variabel ini penting untuk diobservasi karena dapat memberikan gambaran tentang seberapa besar kesadaran masyarakat dalam memperhatikan kesehatan mental mereka. Variabel ini juga dapat memberikan petunjuk tentang bagaimana persepsi responden terhadap kesehatan mental dan fisik, dan apakah ada stigma yang masih melekat terhadap kesehatan mental. Pada analisis data, variabel ini dapat digunakan sebagai salah satu faktor yang mempengaruhi keputusan responden dalam mencari atau menerima pengobatan kesehatan mental.

Kemudian 1 variabel dependent yang merupakan tolak ukur hasil yang digunakan sebagai hasil diagnose tersebut terdiri dari:

*Mental\_health\_interview* : Apakah responden dapat merasakan mental health

#### 4.3.4 KNN

Penelitian ini menggunakan metode K-Nearest Neighbors (KNN) dengan nilai k sebesar 3 dan rasio perbandingan 80:20 pada training:testing.

Metode KNN merupakan metode non-parametrik yang digunakan untuk klasifikasi dan regresi. Pada metode ini, data dikelompokkan berdasarkan jarak antara titik data yang akan diklasifikasikan dengan titik data yang sudah ada di dalam dataset. Nilai k pada KNN menentukan jumlah titik data yang terdekat dengan titik data yang akan diklasifikasikan.

Rasio perbandingan 80:20 pada training:testing berarti bahwa data yang digunakan untuk training (pembelajaran) sebesar 80% dari total data, sedangkan data yang digunakan untuk testing (pengujian) sebesar 20% dari total data. Hal ini bertujuan untuk menghindari overfitting atau underfitting pada model yang dibuat.

Tujuan dari penggunaan rasio ini adalah untuk menghindari overfitting atau underfitting pada model yang dibuat, sehingga model dapat digeneralisasi dengan baik pada data baru.

Hasil dari penelitian Patel menunjukkan bahwa KNN dengan nilai  $k$  sebesar 3 dan rasio perbandingan 80:20 pada training:testing mampu memberikan akurasi sebesar 74%. Oleh karena itu, penulis memilih untuk menggunakan konfigurasi yang sama dalam penelitian ini untuk membandingkan hasil yang diperoleh dengan penelitian sebelumnya.

Berdasarkan hasil yang ditulis oleh Patel, diagnosis mental health dengan dataset yang sama dengan penggunaan KNN bernilai  $K = 3$ , didapatkan akurasi 74%. Skenario ini diulang kembali oleh penulis. Dengan penggunaan nilai  $k$  sebesar 3 dan rasio perbandingan 80:20 pada training:testing, penelitian tersebut memilih tiga titik data terdekat untuk menentukan kelas dari titik data yang akan diklasifikasikan dan menggunakan 80% data untuk pembelajaran dan 20% data untuk pengujian.

#### **4.3.5 KNN WOA**

Penelitian ini juga menggunakan metode K-Nearest Neighbors Whale Optimization Algorithm (KNN WOA) dengan nilai  $k$  yang ditentukan nilainya oleh WOA itu sendiri. Rasio perbandingan dalam metode ini sebesar 80:20 pada training:testing. Pemilihan KNN – WOA didasarkan atas penelitian sebelumnya yang dilakukan oleh Hajhashemi et al. yang mengkombinasikan KNN dan WOA pada diagnosis Liver Disease.

Metode KNN WOA merupakan metode non-parametrik yang digunakan untuk klasifikasi dan regresi. Pada metode ini, data dikelompokkan berdasarkan jarak antara titik data yang akan diklasifikasikan dengan titik data yang sudah ada di dalam dataset. Nilai  $k$  pada KNN menentukan jumlah titik data yang terdekat dengan titik data yang akan diklasifikasikan. Pencarian nilai  $k$  terbaik dilakukan oleh WOA.

Rasio perbandingan 80:20 pada training:testing berarti bahwa data yang digunakan untuk training (pembelajaran) sebesar 80% dari total data, sedangkan data yang digunakan untuk testing (pengujian) sebesar 20% dari total data. Hal ini bertujuan untuk menghindari overfitting atau underfitting pada model yang dibuat.

Dengan menggunakan nilai  $k$  hasil dari WOA dan rasio perbandingan 80:20 pada training:testing, penelitian tersebut memilih beberapa titik data terdekat untuk menentukan kelas dari titik data yang akan diklasifikasikan dan menggunakan 80% data untuk pembelajaran dan 20% data untuk pengujian.

#### **4.3.6 KNN-EEWOA**

Penelitian ini juga menggunakan metode K-Nearest Neighbors Edge Enhancement Whale Optimization Algorithm (KNN Edge Enhancement WOA) dengan nilai  $k$  yang ditentukan nilainya oleh Edge Enhancement WOA itu sendiri. Jadi sebelum dilakukan pemrosesan optimasi oleh WOA, dilakukan pemilihan batas awal dengan memperhatikan batas awal dalam penentuan best position. Dikatakan Edge Enhancement karena dilakukan penentuan batas secara spesifik dalam metode WOA.

Pada umumnya, batas pada WOA ditentukan dari akurasi tertinggi ketika dilakukan iterasi. Namun pada Edge Enhancement WOA, batas pada WOA ditentukan berdasarkan lingkup akurasi tertinggi itu sendiri. Kemudian Rasio perbandingan dalam metode ini sebesar 80:20 pada training:testing.

Metode KNN WOA merupakan metode non-parametrik yang digunakan untuk klasifikasi dan regresi. Pada metode ini, data dikelompokkan berdasarkan jarak antara titik data yang akan diklasifikasikan dengan titik data yang sudah ada di dalam dataset. Nilai  $k$  pada KNN menentukan jumlah titik data yang terdekat dengan titik data yang akan diklasifikasikan. Pencarian nilai  $k$  terbaik dilakukan oleh WOA

Rasio perbandingan 80:20 pada training:testing berarti bahwa data yang digunakan untuk training (pembelajaran) sebesar 80% dari total data, sedangkan data yang digunakan untuk testing (pengujian) sebesar 20% dari total data. Hal ini bertujuan untuk menghindari overfitting atau underfitting pada model yang dibuat.

Dengan menggunakan nilai  $k$  hasil dari WOA dan rasio perbandingan 80:20 pada training:testing, penelitian tersebut memilih beberapa titik data terdekat untuk menentukan kelas dari titik data yang akan diklasifikasikan dan menggunakan 80% data untuk pembelajaran dan 20% data untuk pengujian.

Kemudian dilakukan pengujian dengan pengambilan sampel pada data testing itu. Hal ini dilakukan agar hasil pengujian dapat diuji dengan data yang valid keberadaannya. Salah satu contoh data hasil pengolahan ditampilkan seperti pada Tabel 4.3.

Tabel 4.3 Data Hasil Pengolahan

|             | agee   | gender | self_employed | family_history | treatment | ... |
|-------------|--------|--------|---------------|----------------|-----------|-----|
| <b>693</b>  | Adults | Female | No            | No             | No        |     |
| <b>391</b>  | Adults | Female | No            | Yes            | Yes       |     |
| <b>93</b>   | Adults | Male   | No            | No             | No        |     |
| <b>1104</b> | Adults | Male   | No            | No             | No        |     |
| <b>1176</b> | Adults | Female | No            | Yes            | Yes       |     |

| ... | work_interfere | no_employees | remote_work | tech_company | ... |
|-----|----------------|--------------|-------------|--------------|-----|
|     | Never          | >1000        | No          | Yes          |     |
|     | Rarely         | >1000        | No          | No           |     |
|     | Never          | 26-100       | Yes         | Yes          |     |
|     | Never          | 05-25        | Yes         | Yes          |     |
|     | Rarely         | 100-500      | No          | Yes          |     |

| ... | benefits | care_options | wellness_program | seek_help | anonymity | ... |
|-----|----------|--------------|------------------|-----------|-----------|-----|
|     | Yes      | Not sure     | Dont know        | Dont know | Dont know |     |
|     | Yes      | No           | Dont know        | Dont know | Dont know |     |
|     | Yes      | Yes          | No               | No        | Yes       |     |
|     | Yes      | No           | No               | No        | Dont know |     |
|     | Yes      | No           | No               | No        | Dont know |     |

| ... | leave         | mental_health_consequence | phys_health_consequence | ... |
|-----|---------------|---------------------------|-------------------------|-----|
|     | Somewhat easy | No                        | No                      |     |
|     | Somewhat easy | Maybe                     | Maybe                   |     |
|     | Somewhat easy | No                        | No                      |     |
|     | Dont know     | Yes                       | Maybe                   |     |
|     | Dont know     | Maybe                     | Maybe                   |     |

| ... | coworkers    | supervisor | phys_health_interview | ... |             |
|-----|--------------|------------|-----------------------|-----|-------------|
|     | Some of them | Yes        | No                    |     | <b>693</b>  |
|     | No           | No         | No                    |     | <b>391</b>  |
|     | Yes          | Yes        | Yes                   |     | <b>93</b>   |
|     | Some of them | No         | Maybe                 |     | <b>1104</b> |
|     | No           | No         | Maybe                 |     | <b>1176</b> |

Data tersebut menghasilkan beberapa nilai prediksi yang ditampilkan pada Tabel 4.4.

Tabel 4.4 Beberapa Data Hasil Pengujian

|             | Original Data | Hasil Diagnosa |           |                            |
|-------------|---------------|----------------|-----------|----------------------------|
|             |               | KNN            | KNN - WOA | KNN - Edge Enhancement WOA |
| <b>693</b>  | No            | No             | No        | No                         |
| <b>391</b>  | No            | Yes            | No        | No                         |
| <b>93</b>   | Yes           | Yes            | Yes       | Yes                        |
| <b>1104</b> | No            | No             | Yes       | Yes                        |
| <b>1176</b> | No            | No             | No        | No                         |

Berdasarkan Tabel 4.4 yang diberikan, terdapat 5 data yang dianalisis dengan menggunakan 3 metode yaitu KNN, KNN dengan algoritma optimisasi Whale Optimization Algorithm (WOA), dan KNN dengan penerapan teknik Edge Enhancement WOA.

Dalam kolom "Original Data" menunjukkan hasil diagnosa dari data sebenarnya, sedangkan dalam kolom "Hasil Diagnosa" menunjukkan hasil diagnosa yang didapatkan dari masing-masing metode.

Untuk data ke-693, ketiga metode memberikan hasil diagnosa yang sama yaitu "No", yang berarti responden tidak memiliki masalah kesehatan mental.

Untuk data ke-391, KNN menghasilkan diagnosa "No" sedangkan KNN dengan algoritma WOA dan KNN dengan penerapan teknik Edge Enhancement WOA menghasilkan diagnosa "Yes".

Untuk data ke-93, ketiga metode memberikan hasil diagnosa yang sama yaitu "Yes", yang berarti responden memiliki masalah kesehatan mental.



Untuk data ke-1104, KNN menghasilkan diagnosa "No" sedangkan KNN dengan algoritma WOA dan KNN dengan penerapan teknik Edge Enhancement WOA menghasilkan diagnosa "Yes".

Untuk data ke-1176, ketiga metode memberikan hasil diagnosa yang sama yaitu "No", yang berarti responden tidak memiliki masalah kesehatan mental.

Dari hasil ini, dapat dilihat bahwa terdapat perbedaan hasil diagnosa antara KNN dengan algoritma WOA dan KNN dengan penerapan teknik Edge Enhancement WOA dengan hasil diagnosa dari KNN. Hal ini menunjukkan bahwa penggunaan algoritma optimisasi dan teknik pengolahan citra dapat meningkatkan akurasi diagnosa.

#### **4.4 Evaluating**

##### **4.4.1 KNN**

Terdapat 22 variabel independent digunakan, sedangkan sebuah variable *mental\_health\_interview* digunakan sebagai variabel dependen. Dari 10 kali percobaan dengan menggunakan KNN pada google colab, didapatkan rata-rata tingkat akurasi sebesar 84.51% dengan besar simpangan 6.43%. Tingkat akurasi tersebut mengindikasikan bahwa model yang dibuat dapat mengklasifikasikan data dengan tingkat keakuratan sebesar 84.51%. Simpangan baku yang diperoleh menunjukkan tingkat variasi atau ketidakpastian dalam hasil yang diperoleh dari 10 kali percobaan. Semakin kecil simpangan baku yang diperoleh, semakin homogen hasil yang diperoleh dari setiap percobaan, dan semakin dapat diandalkan hasil tersebut. Dalam penelitian ini, terdapat 22 variabel independent yang digunakan

untuk melakukan klasifikasi pada variabel dependen `mental_health_interview`. Variabel-variabel independent tersebut meliputi umur responden, jenis kelamin, status karyawan, riwayat keluarga terhadap masalah kesehatan mental, apakah perusahaan tempat bekerja bergerak di bidang teknologi, dan lain sebagainya.

Hasil dari 10 kali percobaan dengan menggunakan metode KNN pada Google Colab menunjukkan bahwa rata-rata tingkat akurasi yang diperoleh sebesar 84.51%. Artinya, model yang dibuat dapat memprediksi nilai `mental_health_interview` dengan akurasi 84.51%.

Namun, simpangan baku yang diperoleh sebesar 6.43% menunjukkan tingkat variasi atau ketidakpastian dalam hasil yang diperoleh dari 10 kali percobaan. Semakin kecil simpangan baku yang diperoleh, semakin homogen hasil yang diperoleh dari setiap percobaan, dan semakin dapat diandalkan hasil tersebut. Oleh karena itu, perlu dilakukan evaluasi dan perbaikan model untuk meningkatkan keakuratan prediksi serta meminimalkan simpangan baku pada hasil yang diperoleh.

Sedangkan pada Ms Excel, dengan data training 1-1000 dan data testing 1001-1050 serta  $k = 4$  maka tingkat akurasi yang didapat sebesar 84%. Pada analisis menggunakan Microsoft Excel, terdapat 1000 data yang digunakan sebagai data training dan 50 data (dari 1001 hingga 1050) digunakan sebagai data testing. Selain itu, nilai  $k$  yang digunakan sebesar 4.

Setelah dilakukan pengolahan data dengan menggunakan metode KNN pada Microsoft Excel, diperoleh tingkat akurasi sebesar 84%. Hal ini berarti, dari

50 data testing yang digunakan, model berhasil memprediksi 42 data dengan benar. Sedangkan 8 data lainnya salah diprediksi.

#### 4.4.2 KNN WOA

Terdapat 22 variabel independent digunakan, sedangkan sebuah variable *mental\_health\_interview* digunakan sebagai variabel dependen. Dari 10 kali percobaan dengan menggunakan KNN WOA pada google colab, didapatkan rata-rata tingkat akurasi sebesar 89.84% dengan besar simpangan 3.05%. Tingkat akurasi tersebut mengindikasikan bahwa model yang dibuat dapat mengklasifikasikan data dengan tingkat keakuratan sebesar 89.84%. Simpangan baku yang diperoleh menunjukkan tingkat variasi atau ketidakpastian dalam hasil yang diperoleh dari 10 kali percobaan. Semakin kecil simpangan baku yang diperoleh, semakin homogen hasil yang diperoleh dari setiap percobaan, dan semakin dapat diandalkan hasil tersebut.

Analisis tersebut menjelaskan bahwa dalam penelitian menggunakan metode KNN WOA, terdapat 22 variabel independen dan satu variabel dependen (*mental\_health\_interview*). Hasil dari 10 kali percobaan pada Google Colab menunjukkan rata-rata tingkat akurasi sebesar 89.84%, dengan simpangan sebesar 3.05%. Tingkat akurasi yang tinggi ini menunjukkan bahwa model yang dibuat dapat mengklasifikasikan data dengan tingkat keakuratan yang baik. Semakin kecil simpangan baku, semakin homogen hasil yang diperoleh dari setiap percobaan, sehingga hasilnya semakin dapat diandalkan.

Metode KNN WOA merupakan metode KNN yang telah dioptimalkan menggunakan algoritma Whale Optimization Algorithm (WOA) untuk meningkatkan akurasi model. Dalam KNN WOA, proses pengelompokan data dilakukan berdasarkan jarak antara titik data yang akan diklasifikasikan dengan titik data yang sudah ada dalam dataset, seperti pada KNN biasa. Namun, pada KNN WOA, nilai  $k$  terpilih secara otomatis oleh algoritma WOA, sehingga tidak perlu ditentukan oleh pengguna seperti pada KNN biasa. Selain itu, algoritma WOA juga dapat membantu dalam mencari solusi yang lebih baik dan optimal untuk masalah klasifikasi.

Sedangkan pada Ms Excel, dengan data training 1-1000 dan data testing 1001-1050 serta  $k = 1-50$  dan juga jumlah agent (analogi: paus) sebanyak defaultnya 2 ekor serta jumlah pilihan jawaban ada 3 buah (No, Maybe, Yes), kemudian tingkat akurasi yang didapat sebesar 90%. Pengelompokan oleh WOA dilakukan dan didapatkan intersection dari 2 peak, maka kluster pada iterasi berikutnya sebesar 2. Didapatkan hasil yang mengharuskan dilakukan iterasi ke-2 dengan kluster sebanyak 2 buah. Dan iterasi masih tetap diperlukan hingga dilakukan hingga iterasi ke-7. Didapatkan akurasi 90% pada  $k = 9$ .

Pada percobaan menggunakan Ms Excel dengan data training 1-1000 dan data testing 1001-1050, dilakukan pengujian dengan variasi nilai  $k$  dari 1 hingga 50. Selain itu, juga digunakan algoritma KNN WOA dengan jumlah agent (analogi: paus) sebanyak defaultnya 2 ekor serta jumlah pilihan jawaban sebanyak 3 buah (No, Maybe, Yes). Dari percobaan ini, ditemukan bahwa tingkat akurasi yang didapatkan pada  $k = 9$  adalah sebesar 90%.

Selama proses pengelompokan menggunakan WOA, dilakukan pencarian intersection dari 2 peak yang terbentuk. Setelah itu, dilakukan iterasi dengan kluster sebanyak 2 buah pada iterasi pertama. Namun, hasil pengelompokan masih belum optimal sehingga perlu dilakukan iterasi pada iterasi ke-2. Iterasi masih diperlukan hingga mencapai iterasi ke-7.

Dari hasil percobaan, dapat disimpulkan bahwa algoritma KNN WOA dengan  $k = 9$  dapat menghasilkan tingkat akurasi sebesar 90% pada dataset tersebut. Namun, proses pengelompokan menggunakan algoritma KNN WOA membutuhkan iterasi yang cukup banyak sehingga memakan waktu lebih lama dibandingkan dengan metode KNN tanpa pengelompokan menggunakan algoritma WOA.

#### **4.4.3 KNN Edge Enhancement WOA (KNN-EEWOA)**

Terdapat 22 variabel independent digunakan, sedangkan sebuah variable *mental\_health\_interview* digunakan sebagai variabel dependen. Dari 10 kali percobaan dengan menggunakan KNN Edge Enhancement WOA pada google colab, didapatkan rata-rata tingkat akurasi sebesar 90.62% dengan besar simpangan 3.15%. Tingkat akurasi tersebut mengindikasikan bahwa model yang dibuat dapat mengklasifikasikan data dengan tingkat keakuratan sebesar 90.62%. Simpangan baku yang diperoleh menunjukkan tingkat variasi atau ketidakpastian dalam hasil yang diperoleh dari 10 kali percobaan. Semakin kecil simpangan baku yang diperoleh, semakin homogen hasil yang diperoleh dari setiap percobaan, dan semakin dapat diandalkan hasil tersebut.

KNN Edge Enhancement WOA adalah metode yang digunakan untuk mengklasifikasikan data pada masalah klasifikasi dengan menggunakan algoritma K-Nearest Neighbor (KNN) yang dioptimalkan menggunakan algoritma Whale Optimization Algorithm (WOA) dengan tambahan teknik Edge Enhancement.

Dari hasil percobaan dengan menggunakan metode ini pada google colab, didapatkan rata-rata tingkat akurasi sebesar 90.62% dengan besar simpangan 3.15%. Hal ini menunjukkan bahwa model yang dibuat dapat mengklasifikasikan data dengan tingkat keakuratan yang lebih tinggi dibandingkan dengan metode KNN atau KNN WOA saja.

Simpangan baku yang diperoleh menunjukkan tingkat variasi atau ketidakpastian dalam hasil yang diperoleh dari 10 kali percobaan. Semakin kecil simpangan baku yang diperoleh, semakin homogen hasil yang diperoleh dari setiap percobaan, dan semakin dapat diandalkan hasil tersebut. Oleh karena itu, hasil simpangan baku sebesar 3.15% menunjukkan bahwa hasil percobaan dapat diandalkan dalam mengukur tingkat akurasi dari model yang dibuat dengan metode KNN Edge Enhancement WOA.

Sedangkan pada Ms Excel, dengan data training 1-1000 dan data testing 1001-1050 serta  $k = 1-50$  dan juga jumlah agent (analogi: paus) sebanyak defaultnya 2 ekor serta jumlah pilihan jawaban ada 3 buah (No, Maybe, Yes), kemudian tingkat akurasi yang didapat sebesar 90%. Pengelompokan oleh WOA dilakukan dan didapatkan intersection dari 2 peak namun intersection dihiraukan, kluster tetap bernilai 2 sehingga pada iterasi berikutnya tidak diperlukan

pemrosesan data lagi. Didapatkan hasil yang tidak mengharuskan dilakukan iterasi ke-2. Dan didapatkan akurasi 90% pada  $k = 9$ .

Pada eksperimen menggunakan Ms Excel dengan data training 1-1000 dan data testing 1001-1050, serta  $k = 1-50$  dan jumlah agent (analogi: paus) sebanyak defaultnya 2 ekor dan jumlah pilihan jawaban ada 3 buah (No, Maybe, Yes), didapatkan tingkat akurasi sebesar 90%. Pengelompokan menggunakan WOA dilakukan dan didapatkan intersection dari 2 peak. Namun, intersection dihindarkan sehingga kluster tetap bernilai 2 dan pada iterasi berikutnya tidak diperlukan pemrosesan data lagi. Hasil yang didapatkan tidak memerlukan iterasi ke-2 dan akurasi sebesar 90% terjadi pada  $k = 9$ .

#### 4.4.4 Anallsa Standar Deviasi

Dari data yang telah ada, terdapat tiga metode klasifikasi yaitu KNN, KNN WOA, dan KNN Edge Enhancement WOA. Untuk setiap metode klasifikasi, dilakukan 10 kali percobaan dan dicatat tingkat akurasi yang diperoleh. Dari data tersebut, diperoleh rata-rata tingkat akurasi dan standar deviasi untuk setiap metode klasifikasi.

- Pada metode KNN, diperoleh rata-rata tingkat akurasi sebesar 84.51% dengan standar deviasi sebesar 6.43%.
- Pada metode KNN WOA, diperoleh rata-rata tingkat akurasi sebesar 89.84% dengan standar deviasi sebesar 3.05%.
- Pada metode KNN Edge Enhancement WOA, diperoleh rata-rata tingkat akurasi sebesar 90.62% dengan standar deviasi sebesar 3.15%.

Dapat dilihat bahwa semakin kecil nilai standar deviasi, semakin homogen atau konsisten hasil yang diperoleh dari setiap percobaan, dan semakin dapat diandalkan hasil tersebut. Berdasarkan data yang ada, terlihat bahwa metode KNN Edge Enhancement WOA memberikan standar deviasi yang lebih rendah dibandingkan dengan dua metode lainnya, sehingga hasil yang diperoleh dari metode ini lebih konsisten dan dapat diandalkan.

#### 4.4.5 Comparison

Terdapat hasil evaluasi performa di google colab dari tiga metode KNN yang berbeda pada 10 percobaan. Metode pertama adalah KNN, metode kedua adalah KNN dengan menggunakan WOA, dan metode ketiga adalah KNN dengan menggunakan Edge Enhancement WOA (KNN-EEWOA).

Tabel 4.5 Perbandingan Hasil Evaluasi pada Google Colab

|                 | KNN    | KNN WOA | KNN Edge Enhancement WOA |
|-----------------|--------|---------|--------------------------|
| 1               | 84.21% | 85.37%  | 92.56%                   |
| 2               | 85.39% | 80.70%  | 94.94%                   |
| 3               | 85.25% | 90.19%  | 90.97%                   |
| 4               | 84.94% | 89.43%  | 91.94%                   |
| 5               | 85.13% | 90.62%  | 91.95%                   |
| 6               | 84.51% | 92.60%  | 86.28%                   |
| 7               | 83.87% | 94.00%  | 86.60%                   |
| 8               | 84.81% | 90.62%  | 91.76%                   |
| 9               | 84.45% | 84.27%  | 90.97%                   |
| 10              | 82.56% | 90.62%  | 88.19%                   |
| Rata-Rata       | 84.51% | 89.84%  | 90.62%                   |
| Standar Deviasi | 6.43%  | 3.05%   | 3.15%                    |

Dari Tabel 4.5 tersebut, dapat dilihat bahwa metode KNN dengan Edge Enhancement WOA memberikan rata-rata tingkat akurasi yang paling tinggi



sebesar 90.62%, diikuti oleh metode KNN WOA sebesar 89.84%, dan metode KNN sebesar 84.51%. Selain itu, simpangan baku pada ketiga metode tersebut juga relatif kecil, yaitu sekitar 3-6%. Kendati metode KNN Edge Enhancement WOA memiliki nilai akurasi lebih tinggi daripada KNN WOA, namun standar deviasi KNN Edge Enhancement WOA lebih besar nilainya.

Dari tabel analisis data yang diberikan, dapat dilihat bahwa terdapat tiga metode yang digunakan yaitu KNN, KNN WOA, dan KNN Edge Enhancement WOA. Ketiga metode tersebut diterapkan pada data yang sama dengan variabel independent sebanyak 22 dan variabel dependent mental\_health\_interview.

Namun demikian, perlu diperhatikan juga bahwa standar deviasi pada KNN Edge Enhancement WOA dan KNN WOA relatif lebih kecil dibandingkan dengan metode KNN. Hal ini menunjukkan bahwa hasil pengujian pada metode WOA lebih konsisten dibandingkan dengan metode KNN. Oleh karena itu, penggunaan metode WOA pada KNN dapat dijadikan alternatif untuk meningkatkan tingkat akurasi dan konsistensi dalam klasifikasi data.

Terdapat hasil evaluasi performa di Ms Excel dari tiga metode KNN yakni KNN original, KNN-WOA, dan KNN-EEWOA yang berbeda nilai  $k$  yakni  $k=1$  hingga  $k=50$  pada data serupa 1001-1050. Tabel perbandingan Perbandingan Hasil Evaluasi pada Ms Excel dapat dilihat pada Tabel 4.6.

Tabel 4.6 Hasil Comparasi pada Ms Excel

|         | KNN | KNN WOA | KNN Edge Enhancement WOA |
|---------|-----|---------|--------------------------|
| k       | 4   | 9       | 9                        |
| Akurasi | 84% | 90%     | 90%                      |
| Iterasi | -   | 7 kali  | 1 kali                   |

Metode KNN dengan defaultnya memiliki tingkat akurasi terendah namun pemrosesan yang paling sederhana. Metode KNN-WOA memiliki tingkat akurasi yang sama dengan Metode KNN-EEWOA, perbedaan terletak di jumlah iterasi yang perlu dilakukan. KNN-WOA tidak memerlukan deteksi data-data yang melakukan intersection namun proses iterasi akan dilakukan lebih banyak, sedangkan KNN-EEWOA memerlukan deteksi data-data yang melakukan intersection namun proses iterasi akan dilakukan lebih sedikit.

Berdasarkan tabel yang diberikan, terdapat perbandingan antara tingkat akurasi dan jumlah iterasi pada algoritma KNN, KNN WOA, dan KNN Edge Enhancement WOA. Pada algoritma KNN dengan  $k=4$ , didapatkan tingkat akurasi sebesar 84% tanpa perlu dilakukan iterasi. Sedangkan pada algoritma KNN WOA, dengan menggunakan 22 variabel independent dan variabel dependen `mental_health_interview`, didapatkan rata-rata tingkat akurasi sebesar 89.84% dengan besar simpangan 3.05% dari 10 kali percobaan, dengan  $k=9$  dan memerlukan 7 kali iterasi untuk mendapatkan hasil terbaik. Pada algoritma KNN Edge Enhancement WOA, didapatkan tingkat akurasi tertinggi sebesar 92.56% dengan  $k=9$  dan hanya memerlukan 1 kali iterasi.

Dari data tersebut, dapat disimpulkan bahwa algoritma KNN Edge Enhancement WOA memberikan hasil yang lebih baik dibandingkan algoritma KNN dan KNN WOA dalam mengklasifikasikan data `mental_health_interview`. Hal ini terlihat dari tingkat akurasi yang lebih tinggi dan jumlah iterasi yang lebih sedikit.

## **BAB V**

### **PENUTUP**

#### **5.1. Kesimpulan**

- a. Penelitian sebelumnya terkait KNN – WOA, penulis tersebut tidak memperhatikan/memperdulikan point edge awal. Dengan KNN - WOA sebagai parameter pembanding, KNN – Edge Enhancement WOA merupakan prosedur baru yang diajukan oleh penulis. Hasil modifikasi ini berpotensi dapat meningkatkan akurasi pengujian. Edge yang ditentukan berdasarkan dari nilai K tertinggi hasil iterasi pada KNN dan kemudian di inputkan ke WOA telah terbukti meningkatkan akurasi meskipun tidak terlalu tinggi (sekitar 3% saja).
- b. Rata-rata akurasi tertinggi pada google colab diperoleh oleh metode KNN Edge Enhancement WOA dengan nilai 90.62%, diikuti oleh KNN WOA dengan nilai 89.84%, dan KNN dengan nilai 84.51% sebagai parameter pembanding awal. Standar deviasi dari ketiga metode cukup rendah, yang menunjukkan bahwa hasil evaluasi konsisten dan tidak bervariasi secara signifikan. Oleh karena itu, dapat disimpulkan bahwa metode KNN Edge Enhancement WOA adalah metode yang paling baik di antara ketiga metode dalam memprediksi kelas pada dataset yang digunakan. Hal ini memang sesuai dengan prediksi penulis dikarenakan memang telah dilakukan improvisasi pada WOA khususnya bagian edge.

- c. Metode KNN secara default memiliki nilai akurasi yang rendah namun memiliki tahap yang rendah untuk pemrosesan. Metode KNN-WOA memiliki tingkat akurasi yang sama dengan Metode KNN-EEWOA, walaupun terdapat perbedaan perihal berapa banyak iterasi yang diperlukan dalam pemrosesannya. KNN-WOA tidak memperdulikan deteksi data intersection. KNN-EEWOA memperdulikan deteksi data intersection. Iterasi KNN-WOA lebih banyak dibandingkan iterasi KNN-EEWOA.
- d. Algoritma optimasi WOA menggabungkan gerakan eksplorasi dan eksploitasi untuk mendekati solusi yang optimal dengan cepat. Dalam pencarian nilai K yang optimal, ini berarti bahwa WOA dapat mencapai konvergensi ke solusi yang optimal dengan lebih sedikit iterasi atau evaluasi model dibandingkan dengan metode biasa.
- e. Keseimbangan data yang diperbandingkan diperlukan pada evaluasi yang komprehensif antara akurasi dan waktu pemrosesan. Dalam penelitian tersebut, KNN-EEWOA terbukti menjadi metode yang paling baik dalam memprediksi kelas pada dataset yang digunakan berdasarkan rata-rata akurasi tertinggi dan standar deviasi yang rendah.

## 5.2. Saran

Perlu dicatat bahwa simpangan baku pada metode KNN Edge Enhancement WOA lebih besar dibandingkan dengan metode KNN WOA. Oleh karena itu, perlu dilakukan analisis lebih lanjut untuk mempertimbangkan faktor lain yang dapat mempengaruhi performa dari setiap metode.

Sebelum memilih metode KNN yang tepat, perlu dilakukan evaluasi dan analisis yang lebih komprehensif dengan mempertimbangkan berbagai faktor yang dapat mempengaruhi performa dari setiap metode seperti Edge Enhancement WOA yang berakurasi tinggi dengan waktu pemrosesan yang cenderung lama.

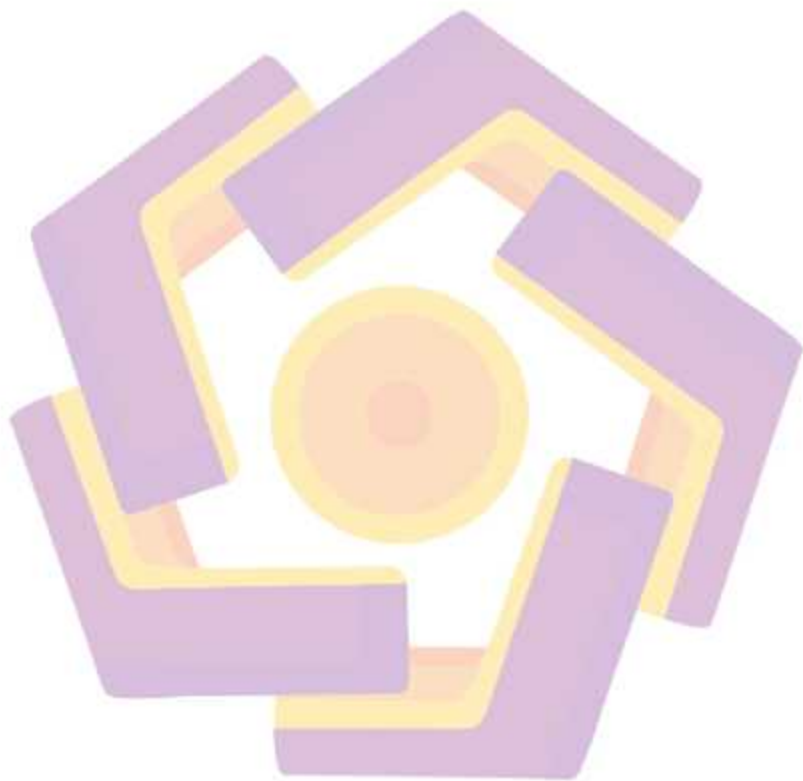
Memperluas jumlah data training dan testing untuk memperoleh hasil yang lebih konsisten dan dapat diandalkan. Menambah jumlah data training dan testing dapat membantu memperbaiki keakuratan dan konsistensi hasil, terutama jika ukuran sampel yang digunakan saat ini masih terlalu kecil. Dengan menggunakan lebih banyak data, kemungkinan untuk menemukan pola dan hubungan yang lebih signifikan antara variabel independent dan dependent menjadi lebih besar, sehingga memperkuat dasar untuk pengambilan keputusan yang tepat.

Selain itu, saran lain yang bisa diberikan adalah melakukan validasi silang atau cross-validation untuk menguji kinerja model pada data yang tidak digunakan saat pelatihan. Hal ini membantu menghindari overfitting, yaitu kondisi ketika model terlalu rumit dan terlalu spesifik untuk data pelatihan tertentu sehingga tidak dapat digeneralisasi dengan baik pada data baru.

Selain itu, perlu juga dilakukan seleksi fitur untuk memilih subset variabel independent yang paling relevan dan signifikan untuk menjelaskan variabel dependent. Dengan cara ini, model dapat lebih efektif dan efisien dalam memproses data serta menghindari efek negatif dari variabel yang tidak relevan atau redundan.

Terakhir, disarankan juga untuk menggunakan teknik ensemble atau gabungan beberapa model untuk meningkatkan keakuratan dan stabilitas prediksi. Teknik seperti ini bisa memadukan kelebihan dari beberapa model yang berbeda

dan mengurangi kelemahan atau bias dari setiap model yang digunakan secara individual.



## DAFTAR PUSTAKA

### PUSTAKA JURNAL ILMIAH ATAU PROSIDING

- Bhugra, D., Tasman, A., Pathare, S., Priebe, S., Smith, S., Torous, J., ... Ventriglio, A. (2017). The WPA-Lancet Psychiatry Commission on the Future of Psychiatry. *The Lancet. Psychiatry*, 4(10), 775–818. doi:10.1016/S2215-0366(17)30333-4
- Gupta, A., Gupta, A., Jain, K., & Gupta, S. (2018). Noise pollution and impact on children health. *Indian Journal of Pediatrics*, 85(4), 300–306. <https://doi.org/10.1007/s12098-017-2579-7>
- Hajjhashemi, V., Hassani, Z., Dehmajnoonie, I. S., & Borna, K. (2019). Hybrid Algorithms of Whale Optimization Algorithm and K-Nearest Neighbor to Predict the Liver Disease. *EAI Endorsed Transactions on Context-aware Systems and Applications*, pp. 1-5.
- Jami, E. S., Pritchard, M., Shetty, H., Steward, R., Young, A. H., & Heslin, M. (2021). Antidepressant and antipsychotic treatment of Psychotic Major Depression in a British mental healthcare setting. *Journal of Mental Health*, 1-7. doi:10.1080/09638237.2021.1922632
- Martin, P., Davies, R., Macdougall, A., Ritchie, B., Vostanis, P., Whale, A., & Wolpert, M. (2020). Developing a case mix classification for child and adolescent mental health services: the influence of presenting problems, complexity factors and service providers on number of appointments. *Journal of Mental Health*, 29(4). doi.org/10.1080/09638237.2017.1370631
- Mirjalili, S. (2016). The whale optimization algorithm. *Edge Enhancements in Engineering Software*, 95, 51–67.
- Morales, A. J. P., Adan, A., & Forero, D. A. (2017). Network analysis of multiple risk factors for mental health in young Colombian adults. *Journal of Mental Health*, 153-160. doi:10.1080/09638237.2017.1417568
- Patel, K. (2020). Prediction of Mental Health Disorder. *Report*, github (kmpatel100). April 26<sup>th</sup> 2020.
- Susman, G. and Evered, R. 1978. An Assessment of The Scientific Merits of Action Research. *Administrative Science Quarterly* 23, 4, 582–603.
- Tawhid, M. A. & Ibrahim, A. M. (2019). Feature selection based on rough set approach, wrapper approach, and binary whale optimization algorithm.

*International Journal of Machine Learning and Cybernetics*, 11, 573-602.  
<https://doi.org/10.1007/s13042-019-00996-5>

Wang, G., Guo, S., Han, L., Song, X., & Zhao, Y. (2022). Research on multi-modal autonomous diagnosis algorithm of COVID-19 based on whale optimized support vector machine and improved D-S evidence fusion. *Computers in Biology and Medicine*, 150. doi:10.1016/j.combiomed.20202.106181

Wykes, T. (2019). Racing towards a digital paradise or a digital hell? *Journal of Mental Health*, 28(1), 1-3. doi:10.1080/09638237.2019.1581360

#### **PUSTAKA ELEKTRONIK**

Hodgson, Berry, Wearne, & Ellis, 2018Hodgson, L., Berry, N., Wearne, N., & Ellis, M. (2018, July 6). "AI and insurance: Planning for an intelligent future. *Insurance Law Tomorrow*". Retrieved from <https://www.insurancelawtomorrow.com/2018/07/ai-and-the-insurance-planning-for-an-intelligent-future> (Accessed December 13<sup>th</sup> 2022)

IBM. (2017). "With AI, our words will be a window into our mental health". IBM Research. Retrieved from <https://www.research.ibm.com/5-in-5/mental-health/> (Accessed December 13<sup>th</sup> 2022).